Copula-based Testing for Dependence Structures

Dominik SZNAJDER

Examination Committee:
Prof. dr. A. Carbonez, chair
Prof. dr. I. Gijbels, promoter
Prof. dr. G. Claeskens, co-promoter
Prof. dr. J. Beirlant
Prof. dr. A.-M. De Meyer
Prof. dr. W. Schoutens
Prof. dr. M. Omelka
(Charles University of Prague,
Czech Republic)
Prof. dr. N. Veraverbeke (UHasselt)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Sciences

October 2011

D/2011/10.705/78
ISBN 978-90-8649-460-6

Acknowledgements

I would like to thank my promoter for giving me the opportunity to discover the world of research and for her guidance. I also wish to thank other researchers for inspiration and encouragement, in particular, my co-promoter and colleagues. Furthermore, I thank the members of my doctoral committee and the examination committee for their valuable comments and remarks. I am also grateful to my family and friends for their support.

Financial support from the GOA/07/04-project “Nonparametric and semi-parametric techniques and robust methods in statistical analysis” of the Research Fund KULeuven and from the IAP research network nr. P6/03 “Statistical Analysis of Association and Dependence in Complex Data” of the Federal Science Policy, Belgium, is gratefully acknowledged.

Abstract in Dutch

In deze thesis gaat de aandacht uit naar toetsingsprocedures voor specifieke afhankelijkheidsstructuren tussen twee stochastische veranderlijken, en in het bijzonder naar kwadrant afhankelijkheid, staart monotoniciteit en stochastische monotoniciteit. Deze types van afhankelijkheid komen vaak voor in verschillende toepassingsgebieden in de wetenschappen en de economie. Als de aanname van een dergelijke specifieke afhankelijkheidsstructuur verantwoord is, dan kunnen meer efficiënte schattingsmethodes worden voorgesteld. De studie van afhankelijkheidsstructuren in een data set verbreedt daarenboven de kennis over de aard van de data en bepalen de belangrijke karakteristieken die in rekening moeten worden gebracht bij het modelleren van associaties.

De thesis bespreekt meerdere data voorbeelden komende uit verschillende toepassingsgebieden, namelijk financiën, verzekeringen, ecologie en micro-economie. De analyses van deze data sets vertonen heel wat gemeenschappelijke kenmerken van afhankelijkheidsstructuren. Vandaar dat in deze thesis een generische marginaal-vrije benadering voor het toetsen van verschillende associaties wordt ontwikkeld.

Het belangrijkste werkmiddel in de toetsingsprocedure is de zogenaamde copula functie. Dit is een bivariate verdelingsfunctie op het eenheidsvierkant met uniforme marginale verdelingen, die de univariate verdelingsfuncties verbindt met hun gezamenlijke verdelingsfunctie. Het bestuderen van specifieke eigenschappen van afhankelijkheden kan dan gebeuren via het bestuderen van de overeenkomstige eigenschappen van de copula functie.

De hoofdaandacht gaat in deze thesis uit naar het niet-parametrisch schatten van een copula functie. Dit laat toe om de beschouwde afhankelijkheidsstructuren op een flexiebele manier te modelleren. Één van de belangrijkste bevindingen en innovatieve resultaten van de thesis is een methode voor het construeren van niet-parametrische copula schatters die aan bepaalde voorwaarden voldoen. Deze methode laat niet alleen toe om tot meer efficiënte schattingsmethodes te komen onder de aanname van spec-

ifieke afhankelijkheidsstructuren, maar verbetert daarenboven de kwaliteit van de overeenkomstige toestingsprocedures voor een specifieke afhankelijkheidsstructuur.

Het toetsen van verschillende afhankelijkheidsstructuren is een eerder onontgonnen terrein in de statistiek en deze thesis levert een hoofdbijdrage hierin. De ontwikkelde toetsen zijn gebaseerd op goed gekozen maten die de afstand beschrijven tussen de niet-parametrische copula schatter en een copula die de specifieke afhankelijkheidsstructuur respecteert.

De statistische besluitvorming is gebaseerd op een benadering van de eindige steekproef verdeling van de toetsingsstatistiek uitgaande van een geschatte copula die aan de opgelegde beperkingsvoorwaarde voldoet. Deze methode levert, in het algemeen, een toetsingsprocedure met een groter onderscheidingsvermogen dan de bestaande (asymptotische) toestingsprocedures.

In deze thesis bestuderen we ook de kwaliteit van de zogenaamde Π-referentie resampling methode en de parametrische (beperkt tot specifieke afhankelijkheid) resampling methode, om tot kritische waarden voor de statistische besluitvorming te komen. Het vergelijken van deze methodes met de niet-parametrische (beperkt tot specifieke afhankelijkheid) resampling methode toont aan dat, zonder een voorkennis over de aard van de data, men er verstandiger aan doet om de niet-parametrische werkwijze te gebruiken omwille van zijn flexibiliteit.

Abstract in English

This thesis describes tests for specific dependence structures between two random variables, in particular: quadrant dependence, tail monotonicity and stochastic monotonicity. These kinds of dependence structures are often encountered in different fields of applications in science and business. If the assumption of a specific dependence structure is justified, then more efficient estimation methods can be proposed. Furthermore, studying dependence structures of a particular data set broadens the knowledge on the nature of the data and indicates their important characteristics that have to be taken into account when modeling associations.

The thesis includes several real data examples coming from fields, e.g., of finance, insurance, ecological studies and micro economics. The analysis of these data sets reveals many common features of dependence structure among these examples. Therefore, a generic marginal-free approach to testing for different associations is developed in the thesis.

The main tool used in the testing procedure is a copula function. It is a bivariate distribution function on the unit square with uniform marginals, which which links the univariate distribution functions and their joint distribution function. Thus, studying particular features of dependence structures can be accomplished by studying the corresponding features of the copula function.

The main emphasis in this thesis is put on the non-parametric estimation methods of a copula function to allow for a flexible way to model the considered dependence structures. One of the main outcomes and innovative results of the thesis is the construction method of the constrained non-parametric copula estimators. Not only does this method allow for the more efficient estimation methods under specific dependence structure assumption, but it also facilitates the performance of the corresponding dependence structure test.

Testing for different dependence structures is an unexplored area in statistics and this is the main contribution of the thesis. The tests are based

on well-chosen measures to describe the distance between the non-parametric copula estimator and a copula respecting the specific dependence structure.

The statistical inference is based on approximated finite sample distribution of the test statistic under the constrained estimated copula distribution. This method yields, in general, higher power in comparison to the existing (asymptotic) methods.

This thesis also investigates the performance of the Π-reference resampling and constrained parametric resampling methods to obtain critical values for statistical decision making. The comparison of these methods with the constrained non-parametric resampling indicates that, without a prior knowledge about the nature of the data, one is much safer when using the non-parametric approach because of its flexibility.

List of abbreviations

C	copula function
$c_u(v)$	$\frac{\partial C(u,v)}{\partial u}$ first order partial derivative of C with respect to u
$PQD(X,Y)$	X and Y are positive quadrant dependent
$NQD(X,Y)$	X and Y are negative quadrant dependent
$LTD(Y\|X)$	Y is left tail decreasing in X
$LTI(Y\|X)$	Y is left tail increasing in X
$RTI(Y\|X)$	Y is right tail increasing in X
$RTD(Y\|X)$	Y is right tail decreasing in X
$SI(Y\|X)$	Y is stochastically increasing in X
$SD(Y\|X)$	Y is stochastically decreasing in X
$\left\{(\hat{U}_i, \hat{V}_i)\right\}$	$\hat{U}_i = \frac{n}{n+1}F_n(X_i)$, $\hat{V}_i = \frac{n}{n+1}G_n(Y_i)$ pseudo-observations
C_n	empirical copula estimator
$\hat{C}_n^{LL}$	kernel local linear estimator of copula C
$\hat{C}_n^{LLS}$	kernel local linear shrunken estimator of copula C
$\hat{C}_n^{MR}$	kernel mirror reflection estimator of copula C
$\hat{C}_n^{MRS}$	kernel mirror reflection shrunken estimator of copula C
KS	Kolmogorov-Smirnov distance
CvM	Cramér-von Mises distance
AD	Anderson-Darling distance

List of papers

This thesis is based on the results of the following papers:

- Gijbels, I., Omelka, M., and Sznajder, D. (2010). Positive quadrant dependence tests for copulas. *The Canadian Journal of Statistics*, 38(4):555–581
- Gijbels, I. and Sznajder, D. (2011a). Positive quadrant dependence testing and constrained copula estimation. Submitted to *The Canadian Journal of Statistics*
- Gijbels, I. and Sznajder, D. (2011b). Testing tail monotonicity by constrained copula estimation. Submitted to *Insurance: Mathematics and Economics.*
- Gijbels, I. and Sznajder, D. (2011c). Testing stochastic monotonicity by constrained copula estimation. Manuscript.

Contents

Chapter 1

Introduction

The study of dependencies in a multivariate distribution setting is of universal interest in a vast majority of modern statistical problems. On the one hand, one can think of descriptive statistics where the interest lies in obtaining aggregate measures of certain dependence relations leading to dependence measures or association measures, e.g., Pearson's correlation coefficient or Kendall's tau. On the other hand, the regression analysis model certain dependence relations between random elements. In other words, one of the key tasks in statistics is to study interactions among random variables and this can be summarized as a general concept of dependence.

This thesis focuses on another approach to study dependence, namely testing for the existence of particular dependence structures. A dependence structure, as understood in the thesis, is a characterization of the joint distribution of a random vector autonomous of its marginal distributions.

The importance of the marginal-free understanding of a dependence structure is depicted in Figure 1.1. There we see the same dependence structure for three different groups of marginal distributions.

An example of a dependence structure is positive quadrant dependence, which means that the probability that two random variables jointly exceed certain levels is greater than the probability that each exceeds the corresponding levels independently. It occurs to be a feature of bivariate distributions that does not depend on the marginal distributions. In other words, it is invariant to probability distribution transformations of the marginals.

A tool that is mainly used in dependence structure modelling and testing is a copula function. It is a multivariate function which links the joint distribution function with the marginal distribution functions. Although, a copula might be defined outside of the probabilistic scope, it is very useful

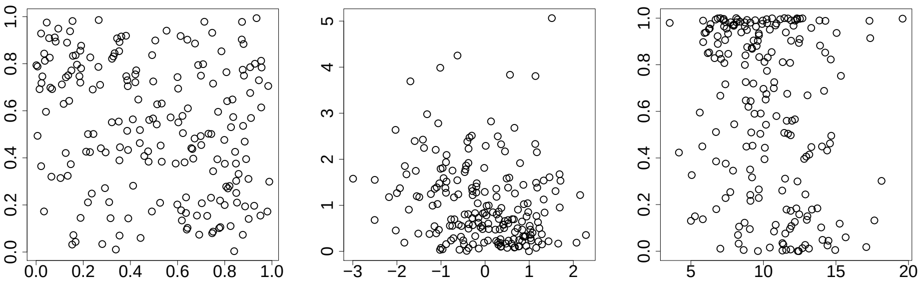


Figure 1.1: *Samples coming from Frank(-1) copula with different marginal distributions.*

to look at it as a multivariate distribution function with uniform margins. A copula function contains complete information about interactions between elements in a random vector. Therefore, it is essentially equivalent to the dependence structure concept. Thus, studying particular features of dependence structures can be accomplished by studying the corresponding features of a copula function, which is specific for a given random vector.

Developing methodology for studying dependence structures is important as the same relation patterns are of interest for various areas of science and business, which are faced with many different kinds of marginal distributions. Although studying dependence structures is broadly explored in dependence modelling, it is not as deeply investigated in testing problems.

The main contributions of this thesis consist of developing statistical tests for specific dependence structures, by exploring its copula characterizations. The methodology is primarily based on finite resampling from semi-parametric and non-parametric constrained copula estimates. The comprised features of dependence structures are quadrant dependence, tail monotonicity and stochastic monotonicity. Quadrant dependence specifically compares the joint distribution function against the independent marginals and can be referred to as positive, negative or neither of the two. It is the weakest form of the considered dependence structures, thus contains the class of tail monotonic structures. Tail monotonicity refers to the left or the right tail and can be increasing, decreasing or neither of them. The strongest dependence structure considered in this thesis is stochastic monotonicity, which can be defined as increasing, decreasing or again neither of the two. Figure 1.2 presents an abstract set containment of the considered characteristics of dependence structures.

Quadrant Dependence
$\cup$
Tail Monotonicity
$\cup$
Stochastic Monotonicity

Figure 1.2: *Abstract inclusions of dependence structures.*

Figure 1.3 depicts a part of one of the real data examples investigated in the next parts of the thesis. It presents insurance claims relating to losses in building value and the profit they generated. With the help of the developed tests in this thesis, one can answer questions such as whether these data reveal a general positive or negative relation, whether there is a monotonic pattern in the corners of the sample plot or whether there possibly is a very strong overall conditional monotonic relation between the two observed variables. The answers to these questions might be influential in the premium setting process or portfolio management for an insurance company.

We shall see that this data set reveals the positive quadrant dependence structure, according to the developed tests. Thus, the next interesting point is to check this positive dependence structure in more detail, namely by looking for existence of any positive tail monotonicity. In this case, the tests strongly reject the hypothesis that the variable 'profits' is left tail decreasing in the variable 'buildings', yet do not reject the hypothesis that the variable 'profits' is right tail increasing in the variable 'buildings'. If right tail increasingness is not rejected, then it is interesting to check further, whether this could be caused by a certain stochastic monotonicity relation.

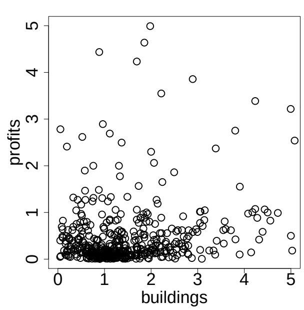


Figure 1.3: *Danish fire insurance data: buildings and profits related claims.*

The next part of the introduction contains a review of the literature concerning copulas in general, examples of its statistical applications in modelling and testing, and estimation techniques. We also briefly review the literature on the discussed dependence structures. We comment on the modern applicability of these dependence structures and on existing competing testing methods. The last part of the introductory chapter will introduce mathematical definitions, properties and notations of the concepts and tools used throughout this thesis.

The following chapters describe tests for positive quadrant dependence (Chapter 2 and Chapter 3, based on papers Gijbels et al. (2010) and Gijbels and Sznajder (2011a)), tests for tail monotonicity (Chapter 4, based on Gijbels and Sznajder (2011b)) and tests for stochastic monotonicity (Chapter 5, based on Gijbels and Sznajder (2011c)). Each chapter also includes a simulation study to investigate the power and size, the finite sample performances of the discussed tests, and to compare these with existing competing methods. Moreover, the tests are applied to a variety of real data examples.

1.1 Brief literature review

The main reference book on the copula theory used throughout this thesis is Nelsen (2006). It contains mathematical definitions and properties of a copula function, methods of construction, links to association measures and specific dependence structures. The crucial theorem on copula decomposition of any multivariate distribution function is thanks to Sklar (1959).

An area where copulas are very frequently used is finance, where they mainly model the co-movement of the financial instruments with the purpose of pricing or risk management. There are several books entirely devoted to these topics, e.g., Cherubini et al. (2004), Xu (2010) and Cherubini et al. (2011) and many devote some chapters to copulas, e.g., in insurance Kaas et al. (2004) or Denuit et al. (2005).

In fact it is hard nowadays to find any domain where there is a need for dependence modelling and where copulas are not taken into account. As an example, there are references to copula usage in modern measurement and advanced psychometrics, e.g., Braeken et al. (2007) and Braeken and Tuerlinckx (2009a,b).

Because of the wide applicability of copulas in dependence modelling, the core focus in testing is being put on goodness-of-fit tests, which check the validity of the applied copula models. Among others see Genest et al. (2006), Genest and Rémillard (2008), Omelka et al. (2009), Berg (2009), Genest et al. (2009b), Genest et al. (2011) and Kojadinovic and Yan (2011).

In terms of estimation of the copula function the first proposed estimator is the empirical estimator of Deheuvels (1979). The maximum likelihood estimator for a copula coming from a parametric family has been described in Genest and Rivest (1993). There has also been considerable research in semi- and non-parametric methods, e.g., Chen and Huang (2007), Omelka et al. (2009) and Genest et al. (2009a). Further developments of the copula fitting problems extended to the concept of a conditional copula are e.g., in Gijbels et al. (2011) and Veraverbeke et al. (2011) and to dynamic stochastic copulas in Hafner and Manner (2010).

The particular dependence structures studied in this thesis originate from Tukey (1958) and Lehmann (1966), and were further developed in Esary and Proschan (1972). Positive quadrant dependence as a testing problem was investigated by Kochar and Gupta (1987) and Janic-Wróblewska et al. (2004), as test for independence against strict positive quadrant dependence. Testing for positive quadrant dependence against not positive quadrant dependence was studied by Denuit and Scaillet (2004) and Scaillet (2005).

Until now, tail monotonicity has not been an object in statistical testing. It has been however explored in the area of positive dependence orderings by Colangelo et al. (2005, 2006) and Colangelo (2008).

When first introduced by Tukey (1958), stochastic monotonicity was referred to as a complete positive/negative regression feature. It is indeed a stronger relation than the nowadays widely discussed problem of (mean) regression monotonicity. In its original form, stochastic monotonicity was tested for by Lee et al. (2009), which also includes an overview of stochastic monotonicity applicability in econometrics.

1.2 Copulas

The word *copula* means from Latin a tie/bond, a friendly/close relationship, according to the William Whitaker's Words Latin dictionary (`http://archives.nd.edu/words.html`).

1.2.1 Definition and basic properties

As already mentioned a copula function has a purely probabilistic interpretation, given in the following definition.

Definition 1.1. *An n-copula function C is any continuous joint distribution function of a random vector of length n, where the marginal distributions are uniformly distributed on the unit interval $\mathbb{I} = [0, 1]$.*

The tool that is mainly used in this thesis is a 2-copula, thus the term 'copula' will refer to that for simplicity. The following theorems are essential for the wide applicability of copulas. The proofs of all the theoretical results in this section can be found in Nelsen (2006) or in the indicated references.

Theorem 1.1 (Sklar (1959))**.**

- *If H is a joint distribution function with marginals $X \sim F$ and $Y \sim G$, then there exists a copula $C_{X,Y}$ (called the copula of X and Y) such that*

  $$H(x, y) = C_{X,Y}\left(F(x), G(y)\right) \quad \forall x, y \in \overline{\mathbb{R}}.$$

  If F and G are continuous, then $C_{X,Y}$ is unique, else $C_{X,Y}$ is uniquely determined on $\mathrm{range}(F) \times \mathrm{range}(G)$*.*

- *If C is a copula and F and G are distribution functions, then the function $H = C(F, G)$ is a joint distribution function, with margins F and G.*

Throughout this thesis we will always assume continuity of the marginal distribution functions.

Furthermore, as long as it is clear from the context, the notation C will be used instead of $C_{X,Y}$ and the marginal distributions of C will be denoted as $U, V \sim \mathcal{U}[0, 1]$, i.e., $(U, V) \sim C$. Moreover, we can interpret the marginals of $C_{X,Y}$ as $U = F(X)$ and $V = G(Y)$.

The following proposition is important for the copula estimation process.

Proposition 1.1. *If H is a joint distribution function with margins F and G, then*

$$C(u,v) = H\left(F^{(-1)}(u), G^{(-1)}(v)\right) \quad \forall (u,v) \in \mathbb{I}^2,$$

where $F^{(-1)}$ and $G^{(-1)}$ are pseudo-inverses of F and G respectively, e.g.,

$$F^{(-1)}(t) = \inf\{x : \ F(x) \geq t\}.$$

Proposition 1.2 is often treated as a definition of a copula function. It defines the boundary (b) and measure (c) conditions of the copula function.

Proposition 1.2. *For each copula C*

(a) $C : \mathbb{I} \times \mathbb{I} \longrightarrow \mathbb{I}$

(b) $C(u,0) = C(0,v) = 0;\ C(u,1) = u;\ C(1,v) = v \quad \forall u,v \in \mathbb{I}$

(c) $C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1) \geq 0 \quad \forall u_1 \leq u_2, v_1 \leq v_2 \in \mathbb{I}.$

Condition (b) outlines the fact that any copula has uniform margins. Condition (c) is a consequence of a copula inducing a probability measure μ_C, building it from rectangles in the unit square $\mathbb{I}^2$, i.e., $\mu_C\big([u_1,u_2] \times [v_1,v_2]\big) = C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1)$, which is also called the C-volume (C-measure) of the rectangle $[u_1,u_2] \times [v_1,v_2]$.

The following theorem is important for the resampling purposes.

Theorem 1.2. *If C is a copula, then*

- *the partial derivative $\frac{\partial C}{\partial u}$ exists for almost all u and for all $v \in \mathbb{I}$, and for such u and v*

$$0 \leq \frac{\partial}{\partial u} C(u,v) \leq 1.$$

- *the partial derivative $\frac{\partial C}{\partial v}$ exists for almost all v and for all $u \in \mathbb{I}$, and for such u and v*

$$0 \leq \frac{\partial}{\partial v} C(u,v) \leq 1.$$

- *The functions*

$$c_v(u) = \frac{\partial C(u,v)}{\partial v} \quad \textit{and} \quad c_u(v) = \frac{\partial C(u,v)}{\partial u}$$

are well-defined and non-decreasing on $\mathbb{I}$. Moreover, they are the conditional distribution functions

$$c_v(u) = \mathbb{P}(U \leq u |\ V = v) \quad \text{and} \quad c_u(v) = \mathbb{P}(V \leq v |\ U = u).$$

A very convenient property of a copula function is its behaviour under monotone transformations of the marginals. One special manifestation of Theorem 1.3 was already mentioned, namely $C_{X,Y} = C_{U,V}$, where $U = F(X)$ and $V = G(Y)$.

Theorem 1.3. *If X and Y are continuous random variables with copula $C_{X,Y}$, then*

- *if α and β are both strictly increasing functions, then*
$$C_{\alpha(X),\beta(Y)}(u,v) = C_{X,Y}(u,v),$$
- *if α is a strictly increasing function and β a strictly decreasing function, then*
$$C_{\alpha(X),\beta(Y)}(u,v) = u - C_{X,Y}(u,1-v),$$
- *if α is a strictly decreasing function and β a strictly increasing function, then*
$$C_{\alpha(X),\beta(Y)}(u,v) = v - C_{X,Y}(1-u,v),$$
- *if α and β are both strictly decreasing functions, then*
$$C_{\alpha(X),\beta(Y)}(u,v) = u + v - 1 + C_{X,Y}(1-u,1-v).$$

Although it is assumed in this thesis that the distributions H, F and G are continuous, this does not mean that H (or C) has a density. We stress it by defining the copula decomposition components in Definition 1.2.

Definition 1.2. *Any copula C can be decomposed in two parts*
$$C(u,v) = A_C(u,v) + S_C(u,v),$$
where
$$A_C(u,v) = \int_0^u \int_0^v \frac{\partial^2}{\partial s \partial t} C(s,t)dtds$$
is called an absolutely continuous component and
$$S_C(u,v) = C(u,v) - A_C(u,v)$$
is called a singular component.

If $C = A_C$, then the copula is said to be absolutely continuous and $\frac{\partial^2}{\partial s \partial t} C(s,t)$ is its joint density. If $C = S_C$, so $\frac{\partial^2}{\partial s \partial t} C(s,t) = 0$ almost everywhere, then the copula is said to be singular.

The support of a copula, defined in Definition 1.3, is closely related to the copula decomposition. In the next section we shall see different copula examples with various decompositions and supports.

Definition 1.3. *The support of the copula is the complement of the union of all open subsets of* $\mathbb{R}^2$ *with* C*-measure zero. If* $\operatorname{supp} C = \mathbb{I}^2$*, then the copula is said to have full support.*

The set of all copula functions is convex (Proposition 1.3) and bounded (Theorem 1.4). Via the convexity property we can obtain interesting copula examples. Bounds provide characteristic limits in the dependence structure, see Definition 1.4.

Proposition 1.3. *A convex linear combination of copulas is a copula.*

Theorem 1.4. *If* C *is a copula, then*

$$\max(u+v-1,0) \le C(u,v) \le \min(u,v) \quad \forall u,v \in \mathbb{I}.$$

Definition 1.4.

- $W(u,v) = \max(u+v-1,0)$ *is called the Fréchet-Hoeffding lower bound,*
- $M(u,v) = \min(u,v)$ *is called the Fréchet-Hoeffding upper bound.*

These bounds in the dependence structure intuitively indicate certain deterministic relations as specified in Proposition 1.4.

Proposition 1.4.

- Y *is an increasing function of* X *almost surely if and only if* $C_{X,Y} = M$*,*
- Y *is a decreasing function of* X *almost surely if and only if* $C_{X,Y} = W$*.*

Furthermore, it can be shown that W and M are valid copula functions themselves. The copulas W and M, which bind the set of all copula functions, are depicted in Figures 1.4 (a) and (c).

The next section gathers other copula examples used throughout this thesis.

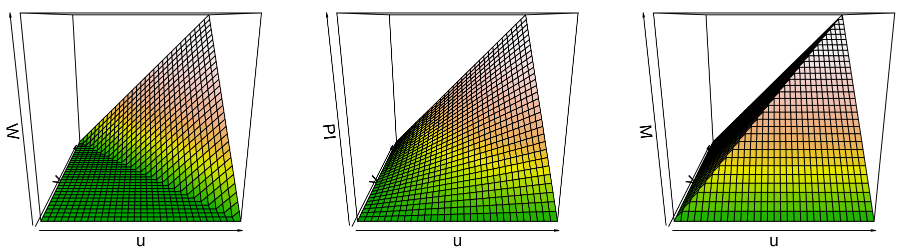
1.2.2 Copula examples

The most characteristic copula is the independence copula, which corresponds to the independence of the random variables. Indeed,

$$\mathbb{P}(X \le x,\ Y \le y) = C\big(F(x), G(y)\big) = F(x)G(y) = \mathbb{P}(X \le x)\mathbb{P}(Y \le y).$$

Definition 1.5. *The independence copula is denoted by* Π*, i.e.,* $\Pi(u, v) = uv$*.*

In Figure 1.4 (b) we depict the independence copula Π.



(a) W (b) Π (c) M

Figure 1.4: *Copulas* W*,* Π *and* M *and corresponding contour plots.*

The rest of the provided examples are families of copulas. The flexibility of a copula family to model the dependence structure can be measured in different ways, yet Definition 1.6 states the nature of what one can call a broad collection of copulas.

Definition 1.6. *If* W*,* Π *and* M *belong to a certain copula family (possibly as the limiting cases), then this copula family is called comprehensive.*

A first copula family to be comprehensive is the Mardia family (Mardia (1970)), which parametrizes a convex mixture of W, Π and M, i.e.,

$$C_{\text{Mardia}} = \frac{\theta^2(1-\theta)}{2} \cdot W + (1-\theta^2) \cdot \Pi + \frac{\theta^2(1+\theta)}{2} \cdot M, \tag{1.1}$$

where $\theta \in [-1, 1]$. An extension of this family is

$$C_{\text{eMardia}} = \omega_W \cdot W + \omega_\Pi \cdot \Pi + \omega_M \cdot M, \tag{1.2}$$

where $\omega_W = \frac{\theta^2(1-\theta)}{2}\gamma$, $\omega_\Pi = (1-\gamma\theta^2)$, $\omega_M = \frac{\theta^2(1+\theta)}{2}$ and $0 < \gamma \leq 1/\theta^2$.

One of the broad collections of copulas is the class of Archimedean copulas, defined in Proposition 1.5.

Proposition 1.5. *If φ is a continuous, convex and strictly decreasing function from $\mathbb{I}$ to $[0,\infty]$ such that $\varphi(1) = 0$, and its pseudo-inverse is defined as*

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{(-1)}(t) & 0 \leq t \leq \varphi(0) \\ 0 & \varphi(0) \leq t \leq \infty, \end{cases}$$

then

$$C(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v))$$

is a valid copula function.

The function φ is called a generator and if $\varphi(0) = \infty$ it is called a strict generator.

Many parametric subclasses arose within the Archimedean copula family, e.g.,

- Frank

  $$\varphi_\theta(t) = -\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}, \quad \theta \in \mathbb{R} \setminus \{0\}$$

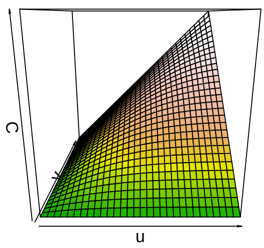
  and $C_{-\infty} = W$, $C_0 = \Pi$ and $C_\infty = M$

- Clayton

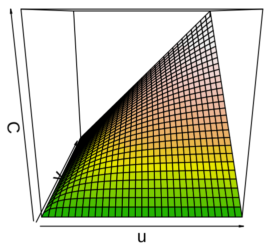
  $$\varphi_\theta(t) = \frac{1}{\theta}\left(t^{-\theta} - 1\right), \quad \theta \in [-1,\infty) \setminus \{0\}$$

  and $C_{-1} = W$, $C_0 = \Pi$ and $C_\infty = M$

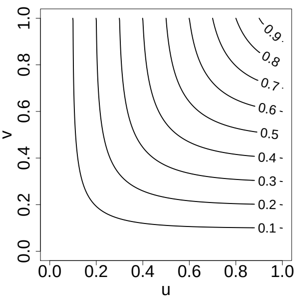
Figures 1.5 (a,b) depict exemplary copula functions coming from these two Archimedean copula families. We can see that they appear to be almost "identical" and, in general, it is extremely hard to distinguish between the copulas on the copula function level. However, if we look at the contour plots of these copula function (c,d) or samples they generate (e,f) we can see that these two copula distributions are very different.



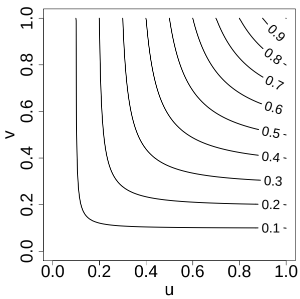
(a) $Frank(5)$



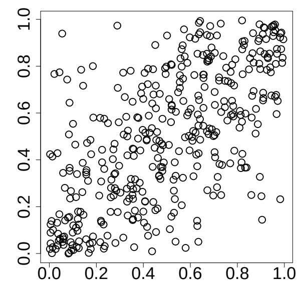
(b) $Clayton(1.745)$



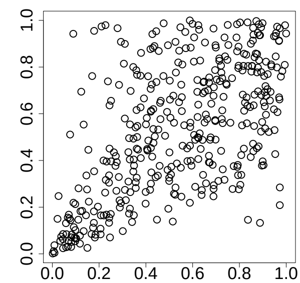
(c) $Frank(5)$



(d) $Clayton(1.745)$



(e) $Frank(5)$



(f) $Clayton(1.745)$

Figure 1.5: *Two examples of Archimedean copula functions together with* 400 *observations from each.*

The last class of copulas to describe in this introduction is a copula obtained by specifying a cross sectional function. Specifically, it is a class of copulas with quadratic sections, which is based on Proposition 1.6.

Proposition 1.6. *If ψ is a function on the unit interval such that*

- *ψ is absolutely continuous on $\mathbb{I}$,*
- *$|\psi'(v)| \leq 1$ almost everywhere on $\mathbb{I}$,*
- *$|\psi(v)| \leq \min(v, 1-v) \quad \forall v \in \mathbb{I}$,*

then

$$C(u,v) = uv + \psi(v)u(1-u)$$

is a valid copula function.

In the next section we recall the main copula estimation method used in the rest of the thesis.

1.2.3 Nonparametric estimation of a copula and resampling

Based on Proposition 1.1 a natural estimator for C can be built on an empirical version of the distribution functions H, F and G. Throughout this thesis we shall use the asymptotically equivalent version of such estimator coming from Deheuvels (1979) and we will refer to this one as the empirical copula estimator, i.e., having a random sample $\{(X_i, Y_i)\}_{i=1}^n$ from (X, Y) then

$$C_n(u,v) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\{\hat{U}_i \leq u,\ \hat{V}_i \leq v\}, \tag{1.3}$$

where $\hat{U}_i = \frac{n}{n+1} F_n(X_i)$ and $\hat{V}_i = \frac{n}{n+1} G_n(Y_i)$, with F_n and G_n the empirical distribution function estimators of X and Y. The values $\hat{U}_i$ and $\hat{V}_i$ are often called the pseudo-observations in the literature.

In Chapter 2 we shall be working with other recently developed nonparametric copula estimators, specifically the kernel local linear estimator of Chen and Huang (2007), the kernel mirror-reflection estimator of Gijbels and Mielniczuk (1990) and their shrunken versions introduced in Omelka et al. (2009). Figure 1.6 gathers the copula estimates corresponding to the samples in Figures 1.5 (c) and (d). As underlined earlier, it is hard to visibly distinguish even very distinctive copula functions, therefore Figure 1.7 presents the corresponding contour plots.

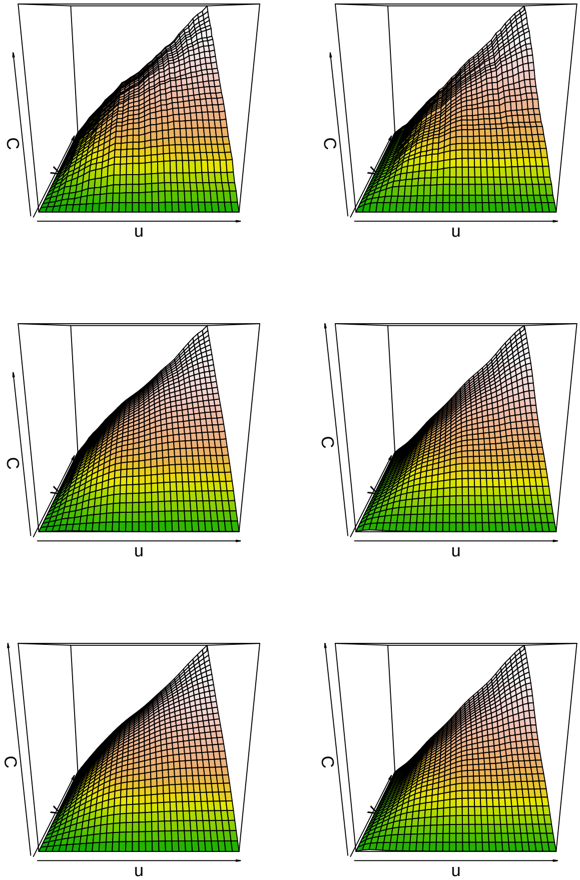


Figure 1.6: *Copula estimates based on the samples in Figure 1.5 (c) (left column) and (d) (right column) for different non-parametric estimators: empirical (top row), kernel local linear shrunken (middle row) and kernel mirror-reflection (bottom row).*

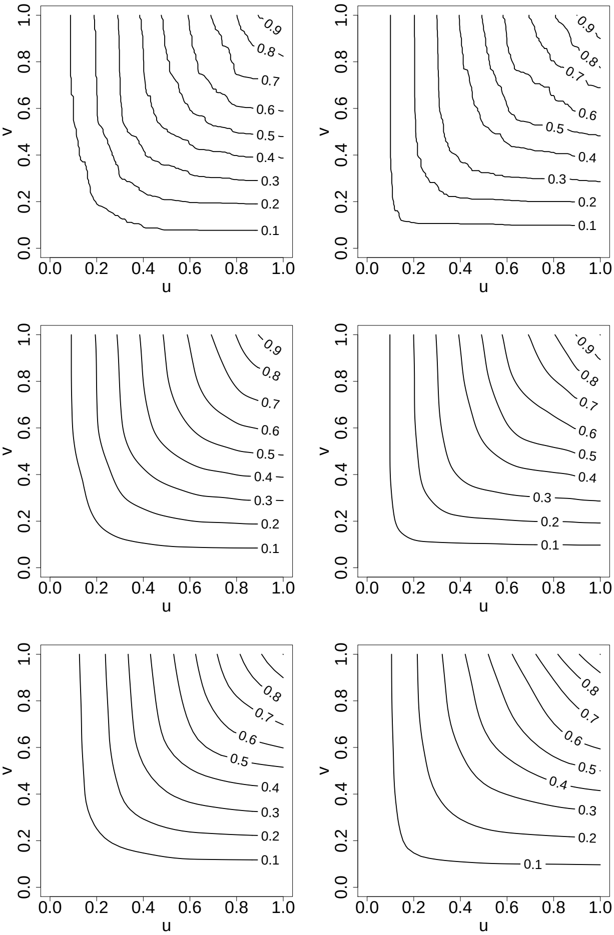


Figure 1.7: *Contour plots for the copula estimates in Figure 1.6.*

It is to be noted that none of these estimators is a valid copula function. In particular, the empirical copula estimator clearly does not fulfill condition (b) in Proposition 1.2 as it is a jump function by construction. All of these estimators are however consistent estimators. Constructing a copula estimator which is a copula itself is not an easy task.

According to Theorem 1.2 we can describe a general resampling process from a given copula C as follows.

Algorithm 1.1.

1. *Draw two observations* u, t *from the uniform distribution on the unit interval*
2. *and compute* $v = c_u^{-1}(t)$.

Now, (u, v) is an observation from the distribution C and we can repeat the process to obtain a sample of observations of a given size.

Having a copula estimator which is itself a copula is essential in resampling. It is possible to approximately resample from the smoothed versions of copula estimators, but it is not enough for the proposed testing procedures. What is needed is a flexible way to resample from a constrained copula estimator. By constrained we mean a copula estimator which (at least approximately) satisfies certain copula shape structure conditions specified in the next section. There is no clear way to modify existing copula estimation methods to obtain a copula estimator constrained to an arbitrary shape. Therefore, a new generic method is proposed in the thesis.

The main concept relies on smoothing the initial discrete constrained copula estimator. Let us specify a grid of points $\{(u_i, v_j)\}_{i,j=0}^{m+1}$ on the unit square and compute the corresponding constrained copula values $c_{i,j}$ for $i, j = 0, \ldots, m+1$. Then we can apply any smoothing technique to obtain the first order partial derivate estimate based on $\{c_{i,j}\}_{i,j=0}^{m+1}$. In this thesis we focus on the local linear smoothing methodology, see e.g., Wand and Jones (1995) and Fan and Gijbels (1996). Specifically, we approximate $c_u(v)$ of the constrained copula estimate in the following way.

$$c_{u,n}(v) = [0, 1, 0](X'WX)^{-1}X'WY, \tag{1.4}$$

where

$$Y = \begin{bmatrix} c_{0,0} \\ \vdots \\ c_{m+1,m+1} \end{bmatrix}, \; X = \begin{bmatrix} 1 & u_0 - u & v_0 - v \\ & \vdots & \\ 1 & u_{m+1} - u & v_{m+1} - v \end{bmatrix},$$

$$W = \begin{bmatrix} k_{h_1}(u_0 - u)k_{h_2}(v_0 - v) & & 0 \\ & \ddots & \\ 0 & & k_{h_1}(u_{m+1} - u)k_{h_2}(v_{m+1} - v) \end{bmatrix}$$

and

$$k_{h_l}(x) = \frac{1}{h_l} k\left(\frac{x}{h_l}\right) \qquad l = 1, 2 \ ,$$

where k is a kernel function (a symmetric probability density function), $h_l > 0$, $l = 1, 2$, are the bandwidth parameters and X' is transpose of X.

To make sure that $c_{u,n}(v)$ is a valid (univariate) distribution function on the unit interval we monotonize it and fit the range properly.

The monotonization technique used throughout the thesis is based on an appealing monotonic rearrangement technique, see e.g., Lieb and Loss (2001).

Definition 1.7. *Let the function ξ denote the increasing rearrangement operator defined as*

$$\xi(f)(t) = \int_A \mathbb{I}(f(x) \leq t)dx \qquad \forall t \in B, \tag{1.5}$$

on the space of univariate real-valued bounded functions $f\colon A \to B$ with A a bounded set.

The idea behind the rearrangement operator is to construct a non-decreasing "inverse" of a given function and apply the operator twice to receive a non-decreasing "version" of the original function. As a result, we obtain a function with appealing properties.

Proposition 1.7. *If ξ is the increasing rearrangement operator of Definition 1.7, then*

(a) ξ is uniquely defined,

(b) $(\xi \circ \xi)(f)$ is a non-decreasing function on A,

(c) if f is non-decreasing, then $(\xi \circ \xi)(f) \equiv f$,

(d) for any non-decreasing function $g\colon A \to B$

$$||(\xi \circ \xi)(f) - g||_{L_p(A)} \leq ||f - g||_{L_p(A)},$$

where $||\cdot||_{L_p(A)}$ denotes the norm of p-integrable functions (on the set A) for $p \geq 1$.

The next section defines the specific dependence structures introduced in the beginning of the chapter.

1.3 Dependence structures

The specific dependence structures that are studied in this thesis are quadrant dependence, tail monotonicity and stochastic monotonicity.

Quadrant dependence is the most general kind of dependence and is closely connected to the common association measures.

1.3.1 Association measures

The relation commonly used in statistics to describe the position of points in a two-dimensional space can be expressed as a relation of the components.

Definition 1.8. *If $(x_i, y_i), (x_j, y_j)$ are two observations from a vector (X, Y) of continuous random variables, then they are concordant if $\{x_i < x_j$ and $y_i < y_j\}$ or $\{x_i > x_j$ and $y_i > y_j\}$, and they are discordant if $\{x_i < x_j$ and $y_i > y_j\}$ or $\{x_i > x_j$ and $y_i < y_j\}$.*

In other words the observations are concordant if $(x_i - x_j)(y_i - y_j) > 0$ and are discordant if $(x_i - x_j)(y_i - y_j) < 0$.

This ordering leads to an aggregate measure of concordance.

Definition 1.9. *If $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ is a random sample of n observations from the vector (X, Y) of continuous random variables and c is the number of concordant pairs and d is a number of discordant pairs, then the Kendall's tau for the sample is*

$$\tau_n = \frac{c-d}{c+d} = \frac{c-d}{\binom{n}{2}}.$$

It is clearly a sample probability of concordance minus sample probability of discordance. On the population level Kendall's tau is defined analogously.

Definition 1.10. *Kendall's tau for the random vector (X, Y) is defined as*

$$\tau_{X,Y} = \mathbb{P}\big((X_1 - X_2)(Y_1 - Y_2) > 0\big) - \mathbb{P}\big((X_1 - X_2)(Y_1 - Y_2) < 0\big),$$

where (X_1, Y_1) and (X_2, Y_2) are independent copies of (X, Y).

It occurs that Kendall's tau does not depend on the marginals and is purely a feature of a dependence structure, thus can be expressed in terms of copulas.

Theorem 1.5. *If X and Y are continuous random variables with copula C, then*

$$\tau_{X,Y} = \tau_C = 4 \iint_{\mathbb{I}^2} C(u,v)dC(u,v) - 1 = 4\,\mathbb{E}\big(C(U,V)\big) - 1. \tag{1.6}$$

For the Archimedean copula family the general expression (1.6) can be rewritten in the following way.

Proposition 1.8. *If C is an Archimedean copula with generator φ, then*

$$\tau_C = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt = 1 - 4\int_0^\infty u\left(\frac{d}{du}\varphi^{[-1]}(u)\right)^2 du.$$

Another commonly used dependence measure is Spearman's rho, which again measures certain concordance discrepancies.

Definition 1.11. *Spearman's rho for the random vector (X,Y) is defined as*

$$\rho_{X,Y} = 3\Big(\mathbb{P}\big((X_1 - X_2)(Y_1 - Y_3) > 0\big) - \mathbb{P}\big((X_1 - X_2)(Y_1 - Y_3) < 0\big)\Big),$$

where (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) are independent copies of (X, Y).

In other words, Spearman's rho measures the concordance and discordance for (X_1, Y_1) and (X_2, Y_3), where X_2 and Y_3 are independent, so their copula is Π. This leads to the following theorem.

Theorem 1.6. *If X and Y are continuous random variables with copula C, then*

$$\rho_{X,Y} = \rho_C = 12\iint_{\mathbb{I}^2} uvdC(u,v) - 3 = 12\iint_{\mathbb{I}^2} C(u,v)dudv - 3.$$

Note that ρ_C can be rewritten as

$$\rho_C = 12\mathbb{E}(U,V) - 3 = \frac{\mathbb{E}(U,V) - \frac{1}{4}}{\frac{1}{12}} = \frac{\mathbb{E}(U,V) - \mathbb{E}U\mathbb{E}V}{\sqrt{\mathrm{Var}(U)}\sqrt{\mathrm{Var}(V)}}$$

and hence Spearman's rho is identical to the Pearson's correlation coefficient for random variables $U = F(X)$ and $V = G(Y)$.

Furthermore, Spearman's rho is a "scaled" volume under the graph of the copula, but as

$$\rho_C = 12\iint_{\mathbb{I}^2} C(u,v)dudv - 3 = 12\iint_{\mathbb{I}^2} \big(C(u,v) - uv\big)dudv, \tag{1.7}$$

it can also be interpreted as proportional to the signed volume between the graphs of the copula C and the product copula Π. Thus ρ_C is a measure of "average distance" between the joint distributions of X and Y (as represented by C) and independence (as represented by Π).

This difference between a given copula C and the independence copula Π is the essence of quadrant dependence.

1.3.2 Quadrant dependence

Definition 1.12. *Random variables X and Y are positively quadrant dependent ($PQD(X,Y)$) if for all $x, y \in \mathbb{R}$*

$$\mathbb{P}(X \le x, Y \le y) \ge \mathbb{P}(X \le x)\mathbb{P}(Y \le y),$$

or equivalently

$$\mathbb{P}(X > x, Y > y) \ge \mathbb{P}(X > x)\mathbb{P}(Y > y). \tag{1.8}$$

By analogy we can define negative quadrant dependence ($NQD(X,Y)$).

In other words, $PQD(X,Y)$ holds if

$$H(x,y) \ge F(x)G(y) \qquad \forall x, y \in \mathbb{R}$$

or equivalently

$$C(u,v) \ge uv \quad \forall u, v \in \mathbb{I}. \tag{1.9}$$

This is thus a feature of the dependence structure regardless of the marginal distributions.

We can now rephrase (1.7) and say that Spearman's rho is a measure of the "average" quadrant dependence of a copula C.

Quadrant dependence is also easily tractable after the monotonic marginal transformations.

Proposition 1.9. *Let X and Y be positively quadrant dependent $PQD(X,Y)$. If α is a strictly increasing function, and β_1 and β_2 are strictly decreasing functions, then*

- *$\beta_1(X)$, $\alpha(Y)$ are negative quadrant dependent $NQD\big(\beta_1(X), \alpha(Y)\big)$,*
- *$\alpha(X)$, $\beta_2(Y)$ are negative quadrant dependent $NQD\big(\alpha(X), \beta_2(Y)\big)$,*
- *$\beta_1(X)$, $\beta_2(Y)$ are positive quadrant dependent $PQD\big(\beta_1(X), \beta_2(Y)\big)$.*

Going back to the given copula examples, we can see that the independence copula can be seen as a boundary point in the sets of both PQD and NQD copulas. This important fact will be explored for testing in the next chapters. As for the other copula examples, we can see that the Mardia families are neither PQD nor NQD, and the quadratic section copula family is PQD if and only if ψ is a non-negative function and NQD if it is non-positive. Positive quadrant dependence in the Archimedean copula class is described in the following proposition.

Proposition 1.10. *If an Archimedean copula C has a strict generator φ, then it is PQD if and only if $-\ln \varphi^{(-1)}$ is subadditive on $(0, \infty)$, i.e.,*

$$\ln \varphi^{(-1)}(x+y) \geq \ln \varphi^{(-1)}(x) + \ln \varphi^{(-1)}(y) \quad \forall x, y \in [0, \infty).$$

In case of the two introduced subclasses of Archimedean copulas, the Frank and Clayton copula families, the PQD feature is preserved for the non-negative parameter values and NQD for the non-positive ones. Figure 1.8 depicts examples of Frank and Clayton copula functions minus the independence copula function Π.

Reformulation of Definition 1.12 evolves in the next studied characteristic of the dependence structure, namely tail monotonicity. We say that X and Y are positive quadrant dependent if

$$\mathbb{P}(Y \leq y | X \leq x) \geq \mathbb{P}(Y \leq y | X \leq \infty). \tag{1.10}$$

Intuitively this means that for any fixed y the conditional probability of $Y \leq y$ given $X \leq x$ is greater than the same probability for $x = \infty$. One way to assure that is to ask this probability to be a decreasing function of x for any fixed y.

1.3.3 Tail monotonicity

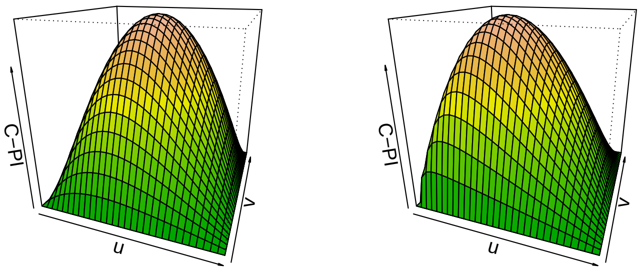
Definition 1.13.

- *Y is left tail decreasing in X ($LTD(Y|X)$) if*

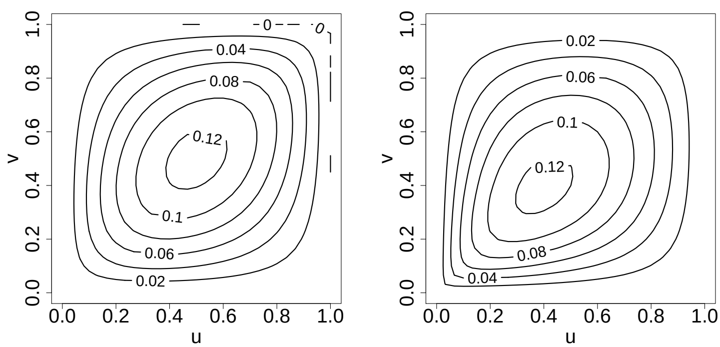
  $$\mathbb{P}(Y \leq y | X \leq x) \quad \textit{is a non-increasing function of } x \textit{ for all } y.$$

- *Y is right tail increasing in X ($RTI(Y|X)$) if*

  $$\mathbb{P}(Y > y | X > x) \quad \textit{is a non-decreasing function of } x \textit{ for all } y.$$



(a) $Frank(5) - \Pi$ (b) $Clayton(1.745) - \Pi$



(c) $Frank(5) - \Pi$ (d) $Clayton(1.745) - \Pi$

Figure 1.8: *Difference between Frank and Clayton copula functions and the independence copula function and the corresponding contour plots.*

By analogy we define left tail increasing and right tail decreasing relations, and also the analogous $X|Y$ relations.

It follows from (1.10) that Y being left tail decreasing in X implies X and Y being positively quadrant dependent, yet even more holds.

Proposition 1.11.

- *$LTD(Y|X)$ or $LTD(X|Y)$ implies $PQD(X,Y)$*
- *$RTI(Y|X)$ or $RTI(X|Y)$ implies $PQD(X,Y)$.*

Note that the opposite does not necessarily hold.

The above tail monotonic properties can also be expressed in terms of copulas. The following expressions are of crucial use in this thesis.

Proposition 1.12.

- *$LTD(Y|X)$ if and only if for every v*

$$\frac{C(u,v)}{u} \quad \textit{is non-increasing in } u \tag{1.11}$$

- *$RTI(Y|X)$ if and only if for every v*

$$\frac{1-u-v+C(u,v)}{1-u} \quad \textit{is non-decreasing in } u,$$

or equivalently, if

$$\frac{v-C(u,v)}{1-u} \quad \textit{in non-increasing in } u.$$

Referring back to the copula examples, we can see that the independence copula fulfills all of the tail monotonic structure conditions, thus can be again treated as a boundary point of each set of copulas. The Mardia families have none of the tail monotonic properties and the considered two subclasses of Archimedean copula families are LTD and RTI for non-negative parameter values, and LTI and RTD for non-positive ones. Furthermore, there is a general condition for the Archimedean copulas to be LTD.

Proposition 1.13. *If an Archimedean copula C has a strict generator φ, then it is LTD if and only if $\varphi^{(-1)}$ is completely monotone on $(0,\infty)$, i.e., it is continuous on $(0,\infty)$ and*

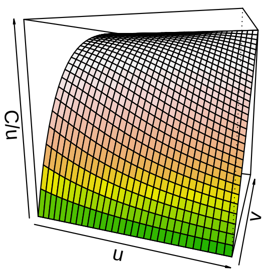
$$(-1)^k \frac{d^k}{dt^k}\varphi^{(-1)}(t) \geq 0 \quad \forall t \in (0,\infty) \quad \textit{and} \quad \mathrm{k}=0,1,\ldots.$$

Moreover, as the Archimedean copula functions are all symmetric in their variables, the tail monotonicity properties are the same for $Y|X$ and $X|Y$.

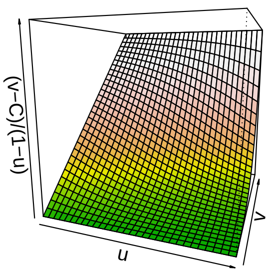
Figure 1.9 presents the functions from Proposition 1.12 C/u and $(v - C)/(1-u)$ reflecting the LTD (Figures 1.9 (a) and (c)) and RTI (Figures 1.9 (b) and (d)) properties of the two examples of Archimedean copula functions.

The cubic section copula family is in general not easily described in terms of tail monotonicity, yet it is easy to see that we have $LTD(V|U)$ and $RTI(V|U)$ if ψ is a non-negative function.

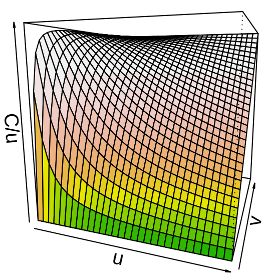
The last considered feature of dependence structure comes from even further narrowing the relation in (1.10).



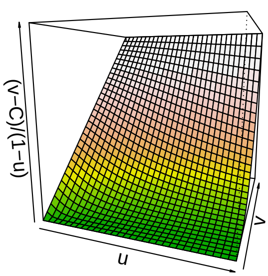
(a) LTD $Frank(5)$



(b) RTI $Frank(5)$



(c) LTD $Clayton(1.745)$



(d) RTI $Clayton(1.745)$

Figure 1.9: *Frank and Clayton functions reflecting LTD and RTI.*

1.3.4 Stochastic monotonicity

Definition 1.14. *Y is stochastically increasing in X ($SI(Y|X)$) if for every y*

$$\mathbb{P}(Y > y|X = x) \quad \textit{is a non-decreasing function of } x.$$

Analogously $SD(Y|X)$ is defined.

Again, note that Y being stochastically increasing in X implies Y being left tail decreasing in X and generalizes to the following.

Proposition 1.14. *$SI(Y|X)$ implies $LTD(Y|X)$ and $RTI(Y|X)$.*

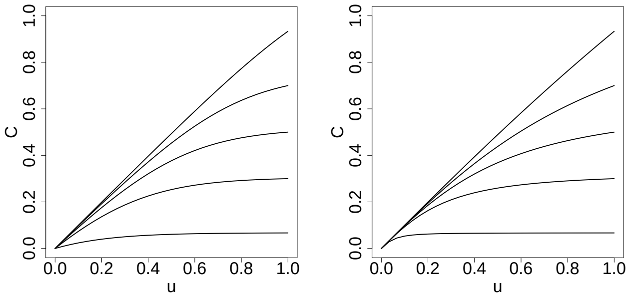
In terms of copulas stochastic monotonicity can be expressed for example in the following way.

Proposition 1.15. *$SI(Y|X)$ if and only if for every v, $C(u, v)$ is a concave function of u.*

For the discussed copula examples the conclusions are the same as for the tail monotonicity case and the general condition for the Archimedean copulas is the following.

Proposition 1.16. *If an Archimedean copula C has a strict generator φ and φ^{-1} is differentiable, then it is SI if and only if $\ln\left(\frac{-d\varphi^{-1}(t)}{dt}\right)$ is convex on $(0, \infty)$.*

From Figure 1.8, where we plot $C - \Pi$, we can deduce that the given examples of Archimedean copulas are concave functions of u for every v. For convenience we also plot several sections of the same copulas in Figure 1.10.



(a) $Frank(5)$ (b) $Clayton(1.745)$

Figure 1.10: *Copula sections $C(u, v)$, for several fixed values of v, for Frank and Clayton copulas.*

Before moving to the next chapter let us refer back to Figure 1.5, where we have seen samples for Frank and Clayton copula family members. The parameters of each were chosen such as to provide the same theoretical value of Spearman's rho equal to 0.64. These samples already look clearly very different from each other. In Figure 1.1 it was only one copula sample transformed with a variety of marginal distributions. For all three joint distributions Spearman's rho equals 0.16. From both figures and examples it is clear that it is hardly possible to “guess” from a scatterplot of the data anything about characteristics of the underlying dependence structure. Therefore, statistical tests that test for a specific dependence structure are needed. This thesis contributes largely to this topic.

Chapter 2

Positive quadrant dependence tests for copulas

2.1 Introduction

This chapter is based on Gijbels et al. (2010) and develops tests for positive quadrant dependence.

A concept that is symmetric to PQD is the concept of NQD, which swaps the inequality in the definition of PQD. The relation between both concepts can be seen in terms of monotonic transformations. Applying increasing functions to X and Y does not change the copula, thus neither their quadrant dependence. However, if an increasing function is applied to one random variable and a decreasing function to the other random variable, then the quadrant dependence of the transformed couple of random variables is changed, see also Proposition 1.9.

Positive quadrant dependence might be a very realistic assumption in many situations. Think of, for example, life expectancies of men and women in various countries. One would expect that a higher life expectancy for men in one country goes along with a higher life expectancy for women in that country. Examples of positive quadrant dependence are ample in particular in insurance and finance. For a discussion about PQD in finance and actuarial sciences see Janic-Wróblewska et al. (2004) and Denuit and Scaillet (2004) and references therein. The knowledge about PQD or NQD of random variables is important for statistical inference. Indeed, if it is reasonable to assume, for example, positive quadrant dependence then such prior knowledge should be exploited in the statistical inference.

Janic-Wróblewska et al. (2004) and Denuit and Scaillet (2004) also investigate testing problems related to this type of dependence structure. In Janic-Wróblewska et al. (2004) rank tests are introduced for testing independence against positive quadrant dependence. Testing for independence against strict PQD was dealt with in Kochar and Gupta (1987). Denuit and Scaillet (2004) test for PQD against non-PQD and construct tests based on a distance concept considering the PQD definitions of both (1.8) and (1.9) using empirical cumulative distribution function estimators.

In this chapter we are concerned with testing the null hypothesis of positive quadrant dependence versus not positive quadrant dependence, focusing as such on finding out whether a PQD assumption is justified. Starting from the PQD characteristic of a copula function given in (1.9), the basic idea of the testing procedures here is to investigate a distance between a non-parametric estimate of the unknown copula and the independence copula function. We consider various non-parametric estimators of a copula function along with three functional distances.

Testing for positive quadrant dependence was also studied in Scaillet (2005). In that paper the author constructs a Kolmogorov-Smirnov type of test based on the empirical copula estimator relying on the asymptotic distribution of the empirical process. Statistical inference is conducted by using a simulation-based multiplier method and a bootstrap method. The present paper contributes further on this testing problem in various aspects. Firstly, testing procedures based on other distance measures such as Cramér-von Mises and Anderson-Darling distance measures (see e.g., Anderson and Darling (1954)) should be studied, since they might reveal different power properties, see also Omelka et al. (2009). Secondly, in recent years other competitive and improved non-parametric estimators of a copula have been introduced and studied and it is worth to investigate how these estimators perform when used in testing procedures. In our study we consider the empirical copula estimator of Deheuvels (1979), kernel type estimators such as the integrated version of the density Mirror Reflection estimator (see Gijbels and Mielniczuk (1990)) and the Local Linear estimator (see Chen and Huang (2007)), as well as recent extensions (improvements) of these two kernel estimators introduced and studied in Omelka et al. (2009). Thirdly, relying on asymptotic theory is not always the best option, since the rate of convergence might require rather large samples before good finite sample behaviour is obtained. We therefore opt for a different approach here, and make use of the independence copula as a reference case included in the null hypothesis. Admittedly this approach also has drawbacks but, as will be seen, these are overruled by the advantages in power performance.

This chapter is organized as follows. In Section 2.2, we briefly discuss the various non-parametric copula estimators and the different test statistics, and establish consistency of these testing procedures. The proofs of these results are given in Section 2.6. Section 2.3 contains a simulation study illustrating the finite sample behaviour of the tests. In Section 2.4 we apply these procedures on real data examples. We conclude in Section 2.5 with some further discussions on the research topic.

2.2 Nonparametric copula estimation and test statistics

Copula estimation is closely related to the estimation of a cumulative distribution function with the main difference that no data from $(F(X), G(Y))$ are observed. Referring to the definition of a copula, an estimation procedure can be divided into two levels, estimation of the marginals and estimation of their joint distribution. If on both levels parametric assumptions are made, then maximum likelihood methods can be applied. However, it is common to make parametric assumptions on the joint level combined with non-parametric estimation of marginals, resulting in popular semi-parametric models. For this usage, there are many well described copula families differing in the number of parameters and characteristics (see e.g., Nelsen (2006)).

In this chapter we are interested in a fully non-parametric approach, and in particular in recently developed estimation procedures described in Omelka et al. (2009). A basic idea behind this and previous estimation methods is to transform the observed data by a monotonic transformation, specifically by the empirical marginal distribution functions, and then to estimate the joint distribution function based on these pseudo-observations. As such we can unify random vectors, which have the same copula, regardless of their marginal distributions.

Suppose we have a sample $(X_1, Y_1), \ldots, (X_n, Y_n) \sim_{iid} H = C(F, G)$. The pseudo-observations as already defined are

$$\hat{U}_i = \frac{n}{n+1} F_n(X_i),\ \hat{V}_i = \frac{n}{n+1} G_n(Y_i),$$

where F_n and G_n are the empirical distributions. The modification $\frac{n}{n+1}$ to the empirical distribution simply pulls the pseudo-observations a bit more away from one (see Genest et al. (1995)). By doing so potential difficulties arising at boundaries can be reduced. The pseudo-observations are then treated as a sample from the random vector $(F(X), G(Y)) \sim C$ and the copula C can be estimated non-parametrically as a bivariate distribution on the unit square. However, because of the unit square domain there are boundary issues arising in the estimation task. Therefore, in our testing procedure, we investigate along with the empirical estimator, the kernel estimators of Chen and Huang (2007) and Gijbels and Mielniczuk (1990) together with their "shrunken" modifications proposed by Omelka et al. (2009) for better consistency results.

In summary our study involves the following copula estimators:

- Empirical copula estimator (Deheuvels (1979))

$$C_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{\hat{U}_i \le u, \hat{V}_i \le v\},$$

  where $\mathbb{I}\{A\}$ denotes the indicator function of A.

- Kernel Local Linear estimator (Chen and Huang (2007)) Denoted by $\hat{C}_n^{\mathrm{LL}}$:

$$\hat{C}_n^{\mathrm{LL}}(u,v) = \frac{1}{n}\sum_{i=1}^{n} K_{u,h_n}\left(\frac{u-\hat{U}_i}{h_n}\right) K_{v,h_n}\left(\frac{v-\hat{V}_i}{h_n}\right),$$

  where h_n is a smoothing parameter and $K_{u,h_n}(x) = \int_{-\infty}^{x} k_{u,h_n}(t)dt$ is the integral of the modified kernel

$$k_{u,h}(x) = \frac{k(x)\left(a_2(u,h) - a_1(u,h)x\right)}{a_0(u,h)a_2(u,h) - a_1^2(u,h)}\mathbb{I}\{\frac{u-1}{h} < x < \frac{u}{h}\},$$

  where

$$a_\ell(u,h) = \int_{\frac{u-1}{h}}^{\frac{u}{h}} t^\ell k(t)dt \quad \text{for } \ell = 0,1,2$$

  and k is a symmetric kernel function that is bounded on the unit interval, e.g., the Epanechnikov kernel $k(x) = 0.75(1-x^2)\mathbb{I}\{|x| \le 1\}$.

- Kernel Local Linear Shrunken estimator (Omelka et al. (2009)) Denoted by $\hat{C}_n^{\mathrm{LLS}}$:

$$\hat{C}_n^{\mathrm{LLS}}(u,v) = \frac{1}{n}\sum_{i=1}^{n} K_{u,h_n}\left(\frac{u-\hat{U}_i}{b(u)h_n}\right) K_{v,h_n}\left(\frac{v-\hat{V}_i}{b(v)h_n}\right),$$

  where $b(w) = \sqrt{\min(w,1-w)}$.

- Kernel Mirror-Reflection estimator (Gijbels and Mielniczuk (1990)) Denoted by $\hat{C}_n^{\mathrm{MR}}$:

$$\hat{C}_n^{\mathrm{MR}}(u,v) = \frac{1}{n}\sum_{i=1}^{n}\sum_{\ell=1}^{9}\left[K\left(\frac{u-\hat{U}_i^{(\ell)}}{h_n}\right) - K\left(\frac{-\hat{U}_i^{(\ell)}}{h_n}\right)\right] \cdot \left[K\left(\frac{v-\hat{V}_i^{(\ell)}}{h_n}\right) - K\left(\frac{-\hat{V}_i^{(\ell)}}{h_n}\right)\right],$$

where $\{(\hat{U}_i^{(\ell)}, \hat{V}_i^{(l)}),\ i=1,\dots,n,\ \ell=1,\dots,9\} =$

$= \{(\pm\hat{U}_i, \pm\hat{V}_i), (\pm\hat{U}_i, 2-\hat{V}_i), (2-\hat{U}_i, \pm\hat{V}_i), (2-\hat{U}_i, 2-\hat{V}_i),\ i=1,\dots,n\}$,

and $K(x) = \int_{-\infty}^{x} k(t)dt$ is the integral of the considered kernel k.

- Kernel Mirror-Reflection Shrunken estimator (Omelka et al. (2009)) Denoted by $\hat{C}_n^{\mathrm{MRS}}$:

$$\hat{C}_n^{\mathrm{MRS}}(u,v) = \frac{1}{n}\sum_{i=1}^{n}\sum_{\ell=1}^{9}\left[K\left(\frac{u-\hat{U}_i^{(\ell)}}{b(u)h_n}\right) - K\left(\frac{-\hat{U}_i^{(\ell)}}{b(u)h_n}\right)\right] \cdot \left[K\left(\frac{v-\hat{V}_i^{(\ell)}}{b(v)h_n}\right) - K\left(\frac{-\hat{V}_i^{(\ell)}}{b(v)h_n}\right)\right].$$

It should be mentioned that in Chen and Huang (2007) the pseudo-observations are obtained via kernel methods. The authors showed however that strong undersmoothing is needed in this step, and hence we here decided directly for a rank estimation, which coincides with the limiting case that the smoothing parameter tends to zero.

The test statistics for testing for positive quadrant dependence are based on distances between the estimated copula and the independence copula. The distances measure the violation part of the copula estimator with the positive quadrant dependence hypothesis under the null. We focus on measures based on L_∞ and L_2 distances. Denote by $\hat{C}_n$ the estimated copula distribution function. We then consider the following statistics

- Kolmogorov-Smirnov

$$S_n^{\mathrm{KS}} = \sqrt{n} \sup_{u,v\in[0,1]} (uv - \hat{C}_n(u,v))_+,$$

where $(\cdot)_+ = \max(\cdot, 0)$

- Cramér-von Mises

$$S_n^{\mathrm{CvM}} = n \int_{\mathbb{I}^2} (uv - \hat{C}_n(u,v))_+^2 d\hat{C}_n(u,v).$$

- Anderson-Darling

$$S_n^{\mathrm{AD}} = n \int_{\mathbb{I}^2} \frac{(uv - \hat{C}_n(u,v))_+^2}{uv(1-u)(1-v)} d\hat{C}_n(u,v) \ .$$

Note that the correction factor $(uv(1-u)(1-v))^{-1}$ in the Anderson-Darling distance puts more attention to the boundaries of a copula. This weight factor is in fact the asymptotic variance of the empirical copula estimator (based on pseudo-observations) when the true underlying copula is the independence copula. Such a weighting factor is also appealing from an intuitive point of view since the closer one gets to the boundaries the smaller the absolute differences between the copulas are. In addition crucial differences between copulas (and therefore between dependency structures) are often hidden close to the boundaries.

With a specified copula estimator and a functional distance, and an i.i.d. sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from a joint distribution with underlying copula C, we can build a test statistic S_n to test the null hypothesis of positive quadrant dependence

$$H_0 : \ \forall u, v \in [0,1] \qquad C(u,v) \geq uv$$

against the negation of this

$$H_1 : \ \exists u, v \in [0,1] \qquad C(u,v) < uv.$$

The distribution of S_n under the null hypothesis is unknown and there are various options to tackle this problem. A first option is to rely on asymptotic results for the copula estimator at hand. For all copula estimators mentioned above weak convergence results are available. Possible drawbacks of this approach are that the asymptotics might kick in only for rather large sample sizes, that the test statistics are non-trivial functionals of the copula estimator, and that it typically requires estimation of partial derivatives of the copula function, resulting in a rather complex estimation procedure. A second approach is to use resampling methods to mimic the distribution of S_n under the null hypothesis. The multiplier and bootstrap methods of Scaillet (2005) follow these approaches. A potential problem with these two methods is that the resampling should in fact be done under the null hypothesis, which cannot be guaranteed.

The approach we follow here consists of making reference to the specific case of the independence copula that is included in the null hypothesis of positive quadrant dependence. To approximate the distribution of S_n under H_0 we draw samples from the independence copula $\Pi(u,v) = uv$, and use these drawings in a Monte Carlo setting to approximate the critical values of the test. Admittedly, this is just selecting one specific element out of the families of all copulas under the null, but the selection makes sense given the importance of the independent case in general.

More precisely, the test works as follows. For a sample of size n with true (unknown) underlying copula C we

$$\text{reject } H_0 \text{ if } S_n > c^{\Pi}_{\alpha,n}, \tag{2.1}$$

where $c^{\Pi}_{\alpha,n}$ is the quantile of the test statistic S_n under the independence copula Π. By using the independence copula to obtain the critical values, we expect to reach the upper bound of the type I error of the test in our composite null hypothesis testing problem. It is also expected that if we move within the null hypothesis set further away from the Π copula (i.e., the larger the discrepancy is between $C(u,v)$ and $\Pi(u,v) = uv$), the smaller the actual significance level of the test will be. This issue is investigated in the finite sample study in Section 3.3.

Equivalent to calculating critical values in hypothesis testing is to calculate the p-value, the probability (under the null) that the considered test statistic exceeds its observed value. In practice this leads to a rejection rule based on an estimated p-value denoted by p_n:

$$\text{reject } H_0 \text{ if } p_n < \alpha.$$

In a bootstrap or multiplier method an estimator for the p-value is

$$p_{n,m} = \frac{1}{m}\sum_{i=1}^{m} \mathbb{I}\{S_{n,m} > S_n\}, \tag{2.2}$$

where $S_{n,m}$ is obtained either from a bootstrap $S^{(\mathrm{B})}_{n,m}$ or a multiplier $S^{(\mathrm{M})}_{n,m}$ method, see Scaillet (2005).

To guarantee the consistency of the proposed testing procedure (2.1) some regularity assumptions are needed. These differ for the different estimators (see Table 2.1).

(C1) The first order partial derivatives of C with respect to u and v, denoted by c_u and c_v respectively, are continuous on the set $[0,1]^2 \setminus \{(0,0),(0,1),(1,0),(1,1)\}$.

(C2) The second order partial derivatives of C denoted by c_{uu}, c_{uv} and c_{vv}, satisfy

$$c_{uu}(u,v) = O\left(\tfrac{1}{u(1-u)}\right), \ c_{vv}(u,v) = O\left(\tfrac{1}{v(1-v)}\right),$$

$$c_{uv}(u,v) = O\left(\tfrac{1}{\sqrt{uv(1-u)(1-v)}}\right).$$

(C3) The second order partial derivatives c_{uu}, c_{uv} and c_{vv} are bounded on $[0,1]^2$.

(C4) Let μ_C be the measure associated with a copula C, λ_2 the Lebesgue measure on $[0,1]^2$, $I_0 = \{(u,v) \in [0,1]^2 : \ C(u,v) = uv\}$ and ∂I_0 the boundary of the set I_0, then $\mu_C(\partial I_0) = \lambda_2(\partial I_0) = 0$.

(Bw) The bandwidth h_n satisfies $h_n = O(n^{-1/3})$.

Table 2.1 lists the assumptions which are needed for the weak convergence of the process $\sqrt{n}(\hat{C}_n - C)$ to a centered Gaussian process G_C in the space of the bounded functions $\ell^\infty([0,1]^2)$, see Theorems 1 and 2 in Omelka et al. (2009).

Estimator	Assumptions
C_n	**C1**
$\hat{C}_n^{\mathrm{LL}}$, $\hat{C}_n^{\mathrm{MR}}$	**C3**, **Bw**
$\hat{C}_n^{\mathrm{LLS}}$, $\hat{C}_n^{\mathrm{MRS}}$	**C1**, **C2**, **Bw**

Table 2.1: *Assumptions for the various estimators.*

The limiting Gaussian process G_C has on $[0,1]^2 \setminus \{(0,0),(0,1),(1,0),(1,1)\}$ the representation

$$G_C(u,v) = B_C(u,v) - c_u(v)\, B_C(u,1) - c_v(u)\, B_C(1,v) \ , \tag{2.3}$$

where B_C is a two-dimensional pinned C-Brownian sheet on $[0,1]^2$, i.e., it is a centered Gaussian process with covariance function

$$\mathrm{E}\left[B_C(u,v)B_C(u',v')\right] = C(u \wedge u', v \wedge v') - C(u,v)C(u',v') \ , \tag{2.4}$$

and G_C is defined to be zero in the "corner points" $\{(0,0),(0,1),(1,0),(1,1)\}$.

Note that from Condition **(C1)** (or also from Condition **(C3)**) and the properties of a copula it follows that the limiting process G_C is equal to zero on the boundaries of the unit square $[0,1]^2$. This issue will be used in the proofs.

Theorem 2.1 below establishes the consistency results for testing procedure (2.1), and is similar to results provided in Scaillet (2005) for the tests therein. The proof of the theorem is given in Section 2.6.

Theorem 2.1. *If $\sqrt{n}(\hat{C}_n - C)$ converges weakly to a zero mean Gaussian process and* **(C4)** *holds, we have that*

(i) if H_0 is true, then

$$\lim_{n\to\infty} \mathbb{P}(\textit{reject } H_0) = \lim_{n\to\infty} \mathbb{P}\left(S_n > c^{\Pi}_{\alpha,n}\right) \leq \alpha,$$

(ii) if H_0 is false, then

$$\lim_{n\to\infty} \mathbb{P}(\textit{reject } H_0) = 1.$$

It finally should be mentioned that we rely on Monte Carlo simulations to approximate the quantile $c^{\Pi}_{\alpha,n}$. This does not affect the above consistency results, as the Monte Carlo simulation can be carried out with high precision.

Note that condition **(C4)** is in fact not necessary for the limiting results on the Kolmogorov-Smirnov test statistic, as can be seen from the proof in Section 2.6. Moreover, condition **(C4)** can in fact be weakened a bit, by assuming that $\mu_C(B) = \lambda_2(B)$ for all measurable subsets B of I_0. Both conditions, **(C4)** and this weaker version, are very mild conditions, and we were not able to construct examples for which these are violated.

2.3 Simulation study

In this section we investigate the finite sample behaviour of the testing procedure based on the various non-parametric estimators and the different distance measures. The simulation study mainly focuses on evaluations of powers of the tests in Sections 2.3.1 and 2.3.2, but also studies on the actual size of the tests are provided in Section 2.3.3. It also includes a comparison with the testing procedures of Scaillet (2005) and moreover extends this paper with results on bootstrap and multiplier based testing methods for other distance measures.

Since the tests are based on non-parametrically estimating the marginal distributions using ranks, they are unaffected by monotonic transformations.

Therefore, for the purpose of the study, a direct sampling from the concerned copula distribution was done. For the multiplier method the samples were additionally transformed to have exponential marginal distributions with parameter 1. For all computations the R software (see R Development Core Team (2011)) was used, in particular the copula package of Yan (2007) and Kojadinovic and Yan (2010).

Throughout the simulation study the significance level is 0.05 and the sample size is 200. Computation of the critical values was based on 10000 samples from the independence copula and the power performance was based on 1000 samples. For the Kolmogorov-Smirnov based test statistics, the supremum is searched for on an equally spaced grid of points with distance 0.05 between two consecutive grid points. For the bootstrap and the multiplier methods, we used 1000 repetitions for approximating the p-value, as given in (2.2).

Several copula models are considered in the simulation study. These also include models studied in Scaillet (2005), namely Frank, Gaussian and Farlie-Gumbel-Morgenstern (FGM) copula families for Kendall's tau equal to -0.11 and -0.16. In order to have some more challenging testing problems we also discuss results for two families of mixtures of copulas. A mixture of Frank copulas introduces different dependency structure than these of the frequently used copula families. A second family of mixtures of copulas is the extended Mardia family, which is a convex mixture of the Frechét-Hoeffding boundary copulas and the independence copula.

2.3.1 Classical copula families

Simulations were done for five classical copula families: Frank, Clayton, Gumbel, Gaussian and FGM copulas; and this for two different values of Kendall's tau: -0.11 and -0.16. The Clayton and Gumbel copula families are often used for modeling heavy dependencies in right tails. However, all members of the Gumbel family have the positive quadrant dependence property. Therefore, this family cannot be directly included in the power study, as there are no members violating the PQD condition. It is however interesting to investigate the power of the tests for copulas for which there is a heavy negative quadrant dependence. Transferring the heavy tails from the upper-right corner to the bottom-right corner can easily be obtained by considering $(U, 1-V)$, where $(U, V) \sim C_{\text{Gumbel}}$. This construction preserves the absolute magnitude of Kendall's tau, but changes the sign, so in the above sense we refer to this Gumbel copula as a Gumbel copula with a negative tau.

Table 2.2 presents the simulation results on the power study for the five classical copula families (with significance level 0.05). The entries ‘M’ and ‘B’ in the tables are the results obtained by the Multiplier and Bootstrap methods for the empirical copula introduced in Scaillet (2005). It is clear that with the decrease of Kendall’s tau the overall power increases and also the differences between the various tests decrease. Further it is to be noted that among the five non-parametric copula estimators and the three distance measures, the worst results are almost always for the testing procedure using the empirical copula estimator and the Kolmogorov-Smirnov distance. The bootstrap method of Scaillet (2005) (and its extension to other distances) works worse than the empirical method (short for the method using the empirical copula with resampling from the independence copula) in case of the Kolmogorov-Smirnov distance, but it works slightly better in cases of the Cramér-von Mises and Anderson-Darling distances. The multiplier method of Scaillet (2005) (and its extensions) works similar to the empirical method for the Kolmogorov-Smirnov distance, but much worse in case of the Cramér-von Mises and Anderson-Darling distances.

The performance results for the Cramér-von Mises and the Anderson-Darling based statistics are more comparable. It seems that for “well behaving” copulas like Frank, Gaussian and FGM the Cramér-von Mises based statistics seem to be working better, but for “heavier-tailed” copulas like Gumbel and Clayton the Anderson-Darling based test statistics perform best.

As for the cross-estimators analysis (comparing the performances of the 5 estimators), the only visible pattern is in the Gumbel and Clayton case, where the mirror type of estimator performs not very well regardless of the distance. Also the local linear estimator combined with the Anderson-Darling distance seems to perform worse in case of the “well behaving” copulas.

In conclusion, overall it seems recommendable to use the mirror reflection shrunken or local linear shrunken estimators combined with the Anderson-Darling or Cramér-von-Mises distance measure.

	Frank			Gaussian			FGM			Gumbel			Clayton		
	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD
							$\tau = -0.11$								
E	.613	.729	.719	.526	.673	.699	.605	.726	.749	.553	.710	.773	.573	.704	.770
LL	.718	.742	.543	.670	.721	.622	.727	.762	.536	.710	.765	.791	.737	.768	.773
MR	.744	.743	.745	.680	.683	.688	.741	.741	.742	.681	.698	.693	.700	.690	.697
LLS	.675	.747	.696	.638	.704	.715	.689	.745	.719	.650	.736	.803	.678	.735	.815
MRS	.675	.745	.738	.638	.701	.723	.689	.745	.765	.650	.728	.783	.676	.730	.787
B	.451	.752	.746	.354	.698	.721	.433	.751	.765	.386	.715	.780	.388	.714	.778
M	.615	.573	.469	.529	.521	.471	.605	.569	.460	.557	.540	.522	.576	.560	.537
							$\tau = -0.16$								
E	.896	.947	.950	.855	.946	.961	.909	.961	.960	.849	.934	.958	.875	.948	.969
LL	.945	.958	.873	.939	.969	.937	.966	.974	.869	.937	.961	.957	.962	.967	.966
MR	.951	.956	.956	.949	.957	.956	.970	.973	.973	.929	.932	.933	.941	.933	.934
LLS	.930	.960	.949	.915	.961	.972	.945	.972	.962	.907	.948	.967	.946	.959	.979
MRS	.930	.959	.955	.915	.960	.970	.945	.971	.970	.906	.946	.965	.944	.957	.970
B	.794	.952	.960	.728	.953	.967	.790	.972	.971	.718	.934	.960	.752	.951	.972
M	.896	.892	.845	.859	.886	.868	.911	.917	.864	.845	.864	.862	.871	.893	.897

Table 2.2: *Power study results for Frank, Gaussian and FGM copulas with Kendall's tau equal to* -0.11 *and* -0.16.

2.3.2 Mixed copulas examples

The copulas in the previous section violated the PQD condition in a simple manner by simply being negative quadrant dependent. In particular this implies that the whole copula function is below the independence copula, and hence the violation is on the whole unit square. However PQDness is a global feature, and it is interesting to see how the tests work on examples where the PQD condition is only locally violated. Therefore we consider in this section two examples of copulas which are neither PQD nor NQD, in contrast to these in the previous analysis.

When it comes to copulas which are only locally NQD, then Kendall's tau is no longer an appropriate measure for the difficulty of a testing problem, especially when a considered copula is on average symmetric around the independence copula. In order to have some guidance regarding the difficulty of a testing problem, there is a need for a different appropriate measure of departure from PQDness.

Two such straightforward measures are a maximum violation measure and a mean of violation measure defined respectively as

$$a = \max_{u,v} \left(uv - C(u,v)\right)_+ \qquad \text{and} \qquad b = \int \left(uv - C(u,v)\right)_+ dudv.$$

These measures are similar in spirit to some dependence measures discussed in Nelsen (2006). We calculated the measures a and b for the copulas considered in Section 2.3.1. For all these copulas the values of a and b for specific values of Kendall's tau turned out to be very similar. Table 2.3 presents approximate average values with respect to tau.

τ	a	b
-0.11	0.029	0.014
-0.16	0.043	0.020

Table 2.3: *Approximate average values of a and b for Frank, Gaussian, FGM, Gumbel and Clayton copulas with different taus.*

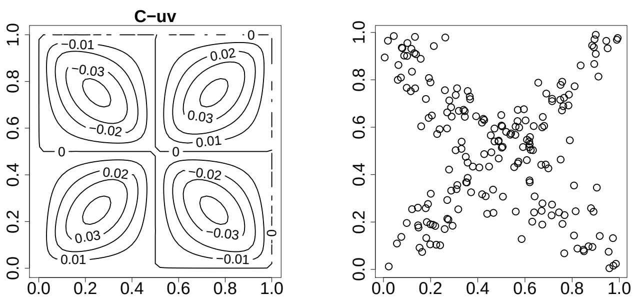
We now consider copulas that are convex mixtures of other copulas and violate the PQDness condition only locally instead of globally.

A first mixture that we study is a symmetric mixture of Frank copulas

$$(U,V) \sim 0.5 \cdot C_{\text{Frank}(\theta)} + 0.5 \cdot C_{\text{Frank}(-\theta)}. \tag{2.5}$$

To give some insight in such a mixture of copulas, we present in Figure 2.1 (a) a contour plot of the difference between such a mixture copula (with

$\theta = 17.5$) and the independence copula. Figure 2.1 (b) depicts a sample from such a copula. Note that for increasing θ one moves further away from the independence case. The larger θ becomes the more concentrated the observations are along both diagonals. Table 2.4 provides the values of a and b for a set of parameters theta.



(a) (b)

Figure 2.1: *(a). Contour plot of the difference between the convex mixture of Frank copulas given by (2.5) and the independence copula; (b) a sample from this convex mixture distribution.*

θ	a	b
1.47	0.002	0.0003
9.5	0.029	0.0060
17.5	0.043	0.0084

Table 2.4: *Approximate values of a and b for mixture of Frank copula with different thetas.*

The parameter values $\theta = 9.5$ and $\theta = 17.5$ were chosen such that the values of a are approximately equal to these from the previous models in Section 2.3.1. The parameter value $\theta = 1.47$ serves as a reference to the case of a single Frank copula with the same parameter (when τ was equal to -0.16). Note from comparing with Table 2.3 that the corresponding b-values are much smaller than these for the "classical" copulas. As will be seen from the simulation results, these mixture copulas present a situation of local violation of the PQD condition that is more difficult to detect. As such we believe that the 'b' measure of PQD-"badness" of a copula is more appropriate than the 'a'-measure.

The simulation results for the mixture of Frank copulas are given in

Table 2.5. The power results for $\theta = 1.47$ are close to the significance levels, which suggests that the mixture in this case is hard to distinguish from the independence case. More informative are the simulation results for $\theta = 9.5$ and $\theta = 17.5$. Although there is much more variability in the simulated rejection probabilities than in the previous study, we again notice that the mirror type of estimator does not perform well regardless of the distance, which was also the case for the Gumbel and Clayton copulas. Regarding the cross-distance analysis it can be concluded, also from these simulation results, that the tests based on the Anderson-Darling distance seem to perform best. We can also see that, when compared to the empirical method, both the bootstrap and the multiplier methods work worse in all of the cases with exception of the multiplier method combined with the Kolmogorov-Smirnov distance.

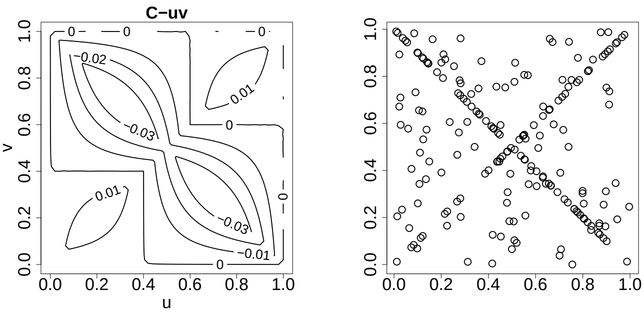
	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD
		θ =1.47			θ=9.5			θ =17.5	
E	.043	.050	.054	.377	.389	.617	.759	.777	.926
LL	.048	.044	.064	.494	.261	.802	.840	.520	.969
MR	.050	.050	.049	.094	.108	.106	.152	.153	.163
LLS	.049	.050	.052	.494	.276	.571	.809	.578	.880
MRS	.049	.050	.048	.491	.269	.527	.810	.561	.853
B	.017	.055	.055	.230	.291	.496	.605	.639	.872
M	.041	.016	.007	.429	.143	.194	.828	.455	.594

Table 2.5: *Simulated rejection probabilities for the mixture of Frank copulas given by (2.5) with parameters* $\theta = 1.47$, 9.5, 17.5.

The second example of models is a mixture of the Fréchet-Hoeffding boundary copulas and the independence copula as defined in (1.2). Rearranging the components in (1.2) yields

$$C = \Pi + \gamma \cdot (C_{\text{Mardia}} - \Pi), \tag{2.6}$$

where C_{Mardia} as defined in (1.1) is the original Mardia copula family. Expression (2.6) reveals the motivation behind this mixture copula – to scale the differences between the Mardia copula and the independence copula. Scaling of this difference does not change the area in the unit square where it is negative or positive. For θ equal to -0.5, -0.2, 0.2, 0.5 respectively there is 0.125, 0.32, 0.68, 0.875 percentage of the area above the independence copula function. For comparison the symmetric mixture of Frank copulas displayed in Figure 2.1 (a) is for half of the area above the independence copula.



(a) (b)

Figure 2.2: *Contour plot of the difference between the extended Mardia copula given by (1.2) with* (θ, γ) *equal to* $(-0.2, 11.9444)$ *and the independence copula function (in (a)) together with a sample from this distribution in (b).*

To give some idea of the form of such copulas, we depict a contour plot of a copula from this family minus the independence copula in Figure 2.2 (a) together with a sample generated from it in Figure 2.2 (b). The parameter θ in this copula is responsible for the concentration of the observations on the diagonals. If θ is positive then it is more probable to have observations from the M copula, which concentrates on the $[(0,0),(1,1)]$ diagonal. If θ is negative then more observations come from the W copula, which concentrates on the $[(0,1),(1,0)]$ diagonal. The parameter γ additionally decreases ($\gamma > 1$) or increases ($\gamma < 1$) the probability of having independent observations in the sample.

Table 2.6 gives values for the a and b measures for the Mardia copula with different parameter values theta. Note that there is no monotonic relation between a and θ.

θ	a	b
-0.5	0.035	0.0103
-0.2	0.004	0.0008
0.2	0.002	0.0002
0.5	0.004	0.0001

Table 2.6: *Approximate values of* a *and* b *for the Mardia copula for different theta values.*

Table 2.9 summarizes the obtained simulation results for some Mardia copulas. Note that the cases with $\theta = -0.2$, 0.2 and 0.5 really represent difficult to very difficult testing problems. Despite the 0.125 area violation of PQDness, the deviation from the independence copula is simply too small for the Mardia copula with parameter 0.5 to be caught by the tests. For $\theta = 0.2$, with the area of violation being 0.32, the powers are mostly below the significance level. Yet, these results are not surprising, when compared to the ones for parameter -0.2. It occurs that the Mardia copula with $\theta = -0.2$ has similar values of a and b as the mixture of Frank copulas with parameter 1.47, which was hard to be distinguished from the independence case. This can be Tables 2.4 and 2.6. The area of violation for the Mardia copula with $\theta = -0.2$ is 0.68, which is more than the 0.5 area of violation for the Frank mixture with parameter 1.47. Thus, the results are expected to be slightly better for the former case. Comparing Tables 2.5 and 2.9 this indeed seems to be the case. Our tests are thus also more sensitive to the case of asymmetric copulas (around the independence copula). The results for the Mardia copula with parameter $\theta = -0.5$ can be interpreted easier after the next simulation part, which incorporates also different values for the γ parameter in (1.2) and (2.6). We can already notice however that in general the bootstrap and multiplier methods work worse than the empirical one, again with exception of the multiplier method combined with the Kolmogorov-Smirnov distance.

Overall the results are better for tests based on the Anderson-Darling distance and the Cramér-von-Mises distance, combined with the shrunken type of kernel estimators.

Thanks to the construction of the extended Mardia copula via the scaling factor γ in (2.6), it is possible to "adjust" γ in order to get a testing model for which the values of a and b are close to these of the previous analysis. By doing so we will be able to compare results for these types of mixture copulas, with local violations of the PQD condition, with copulas for which there is a global violation of this condition. Table 2.7 lists possible values of γ, which give either the same a or b values as for the copulas considered in Section 2.3.1. It is not possible to obtain all values of a and b for all theta, because of the constraint $\gamma\theta^2 < 1$.

Table 2.10 contains the simulation results for the extended Mardia copula in (1.2) with $\theta = -0.5$ and various values of γ. In terms of a (b) the results for γ equal to 0.825 (1.358) can be compared with the results from the study in Section 2.3.1 in case $\tau = -0.11$. The conclusions in terms of possible best choices for distance measures and non-parametric estimators remain

γ	a	b
	$\theta = -0.5$	
0.825	0.029	0.009
1.223	0.043	0.013
1.358	0.048	0.014
1.940	0.068	0.020

γ	a	b
	$\theta = -0.2$	
8.056	0.029	0.006
11.944	0.043	0.010
17.560	0.063	0.014
	$\theta = 0.2$	
18.125	0.029	0.003

Table 2.7: *Approximate values of a and b for the extended Mardia given by (1.2) for different values of gamma and theta.*

the same.

	$\gamma = 8.056$			$\gamma = 11.944$			$\gamma = 17.560$		
	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD
E	.396	.542	.710	.742	.897	.965	.978	.999	1
LL	.408	.387	.791	.721	.664	.963	.953	.913	.996
MR	.256	.285	.281	.440	.491	.487	.711	.755	.751
LLS	.427	.396	.657	.745	.712	.903	.965	.946	.992
MRS	.425	.394	.565	.744	.702	.849	.963	.945	.989
B	.241	.449	.632	.549	.796	.913	.920	.986	.996
M	.396	.287	.293	.743	.678	.677	.978	.964	.965

Table 2.8: *Simulated rejection probabilities for the extended Mardia copula given by (1.2) for $\theta = -0.2$ and different gamma values.*

The same analysis has been done for rescaled Mardia copulas with $\theta = -0.2$ and various values for the γ parameter. We can see from the results presented in Table 2.8 that there is the same gradation in test performance as in case of $\theta = -0.5$ and that similar conclusions hold.

	$\theta = -0.5$			$\theta = -0.2$			$\theta = 0.2$			$\theta = 0.5$		
	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD
E	.592	.746	.805	.067	.081	.092	.037	.041	.040	.001	.001	.007
LL	.582	.635	.836	.081	.082	.130	.037	.036	.076	.004	.001	.146
MR	.517	.564	.555	.073	.071	.072	.035	.039	.038	0	0	0
LLS	.610	.644	.770	.082	.075	.101	.039	.040	.051	.001	.001	.041
MRS	.610	.638	.716	.082	.075	.092	.039	.039	.044	.001	0	.002
B	.413	.673	.755	.029	.084	.097	.018	.043	.046	0	0	.003
M	.598	.522	.508	.066	.038	.020	.036	.014	.008	.001	0	0

Table 2.9: *Simulated rejection probabilities for the Mardia copula with parameters $\theta = -0.5, -0.2, 0.2, 0.5$.*

	$\gamma = 0.825$			$\gamma = 1.223$			$\gamma = 1.358$			$\gamma = 1.940$		
	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD
E	.461	.597	.670	.781	.887	.939	.863	.952	.977	.990	.999	.999
LS	.461	.509	.700	.746	.785	.907	.825	.848	.959	.985	.981	.995
MR	.428	.457	.449	.691	.719	.714	.759	.800	.793	.955	.968	.967
LLS	.472	.520	.624	.776	.803	.897	.848	.870	.925	.987	.990	.997
MRS	.472	.515	.564	.775	.800	.856	.849	.870	.917	.987	.990	.998
B	.299	.542	.605	.634	.838	.899	.729	.906	.949	.973	.997	.998
M	.462	.385	.348	.784	.742	.705	.865	.843	.813	.989	.989	.992

Table 2.10: *Simulated rejection probabilities for the extended Mardia copula given by (1.2) for $\theta = -0.5$ and different gamma values.*

2.3.3 Size simulation study for Frank copula

In this section the aim is to investigate the actual size properties of the proposed tests and to compare these also with the actual size results for the testing procedures of Scaillet (2005). We therefore focus on the Frank copula, since this model served for the simulation study of this kind in Scaillet (2005). In particular we are interested to see how the actual size is influenced when the true copula is equal to the independence copula or the true copula has stronger PQD characteristics.

Results from Table 2.11 suggest that the actual sizes of the tests are quite good and that the significance level holds when the copula does not have too strong PQD-characteristics. Note that for the independence based methods in the case of a true independence copula (the case $\tau = 0$) the differences between the significance level and the actual size can be explained by the Monte Carlo simulation error. The general tendency when going deeper into the null hypothesis (i.e., τ taking on positive values) is clearly visible, since the actual size is decreasing. When getting more and more into violation of the null hypothesis (i.e., τ taking on bigger negative values) the power increases. Again here the test based on the local linear estimator combined with the Anderson-Darling distance measure has the worst performance. This was also noticed from previous simulation results, and hence the use of this specific test statistic should be avoided. One can also see that the bootstrap method combined with Kolmogorov-Smirnov distance and the multiplier method combined with Cramér-von Mises or Anderson-Darling do not hold the level. Moreover, neither of the considered tests holds the level for $C \in H_0$, except for independence based tests in the boundary (independence; $\tau = 0$) case.

	$\tau = -0.07$			$\tau = -0.03$			$\tau = 0$			$\tau = 0.03$			$\tau = 0.07$		
	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD	KS	CvM	AD
E	.337	.405	.408	.117	.135	.135	.056	.059	.061	.008	.011	.010	.003	0	0
LL	.396	.420	.263	.134	.139	.084	.055	.060	.049	.017	.011	.031	0	.001	.008
MR	.413	.410	.411	.131	.128	.130	.059	.061	.060	.012	.008	.008	.001	.001	.001
LLS	.377	.412	.385	.134	.134	.119	.055	.061	.058	.011	.010	.013	.001	.001	0
MRS	.377	.408	.422	.134	.134	.136	.055	.061	.063	.011	.010	.010	.001	.001	0
B	.199	.430	.430	.054	.142	.146	.025	.062	.066	.002	.011	.012	0	0	0
M	.333	.263	.176	.113	.065	.043	.057	.029	.014	.006	.002	.001	.003	0	0

Table 2.11: *Simulation results for power and actual size for Frank copula with different values of Kendall's tau.*

2.4 Applications

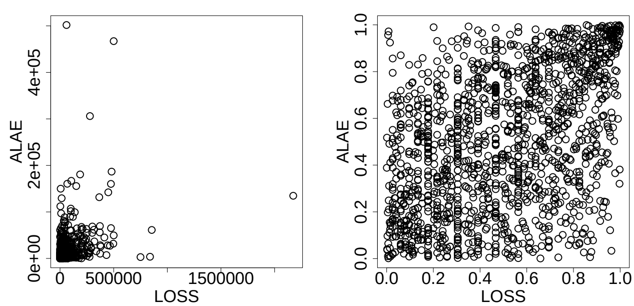
We illustrate the usefulness of the testing procedures on three data examples. A first example concerns the well-known data set from the insurance market introduced by Frees and Valdez (1998) and described in detail in Denuit and Scaillet (2004). The second example is on a data set on life expectancy at birth. This example is similar to the one in Scaillet (2005) but we use an updated data set available from Central Intelligence Agency (2008). Finally, in the last example we investigate the dependence structure between the Belgian stock index BEL20 and the exchange rate between the currencies Euro and American Dollar.

2.4.1 Insurance claim data

This data set consists of 1466 uncensored claims (losses) and claims' costs (ALAE) which are presented together with the corresponding pseudo-observations in Figure 2.3.

The empirical value of Kendall's tau for this sample is 0.31, which is relatively high in comparison with our testing procedure's framework in Section 2.3. In addition the sample size is much higher than the sample size 200 considered in the simulation study. Thus, we expect to observe some positive dependence in this data set.

The approximated p-values provided in Table 2.12 confirm that there is no proof for rejecting the positive quadrant dependency structure between the data in this example.



(a) (b)

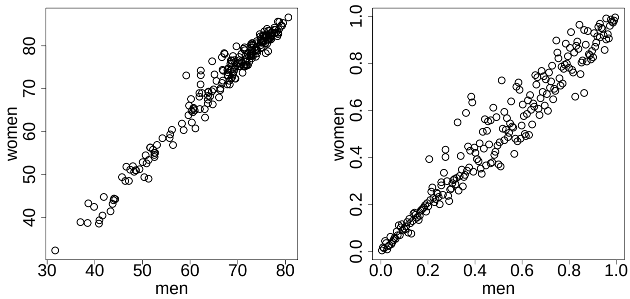
Figure 2.3: *Scatterplot of the insurance claim data (a) and plot of the corresponding pseudo-observations (b).*

	KS	CvM	AD
E	.825	1	1
LL	.989	.990	.975
MR	.812	1	1
LLS	.996	1	1
MRS	.992	.991	.971
B	.989	1	1
M	.813	1	1

Table 2.12: *Approximated p-values for the data set of losses and losses' costs.*

2.4.2 Life expectancy at birth for men and women

This data set consists of estimated life expectancy at birth for men and women in 223 countries. In Figure 2.4 we can see a high concentration of pseudo-observations around the positive diagonal, which suggests that there is strong positive dependence structure. This is also supported by the empirical value of Kendall's tau which equals 0.86. From the approximated p-values in Table 2.13 we can indeed see that there is very strong evidence for not rejecting the positive quadrant dependence. The evidence is even much stronger here than in the previous example, although here the sample size is more than six times smaller.



(a) (b)

Figure 2.4: *Scatterplot of the data concerning life expectancy at birth of men and women (a) and plot of the corresponding pseudo-observations (b).*

	KS	CvM	AD
E	1	1	1
LL	.995	.999	.999
MR	.875	.994	.994
LLS	.999	1	1
MRS	.999	1	1
B	1	1	1
M	1	1	1

Table 2.13: *Approximated P-values for the data set of life expectancy at birth for male vs. female.*

2.4.3 The BEL20 index and the EUR/DOL exchange rate

This data set consists of observations on the Belgian stock index BEL20 and the currency exchange rate EUR/DOL from the period January 2, 2008 till January 8, 2009, resulting into 259 common observations. Figure 2.5 depicts the log-returns of the index and the currency exchange rate together with a plot of their pseudo-observations. In contrast to the previous examples, it is less clear from the plots whether the data are positive quadrant dependent or not.

The empirical Kendall's tau value equals -0.06. This suggests that there is no positive dependence. However, when we look at Table 2.14 then, with 5% significance level, there is no evidence to reject the null hypothesis, except in the case of one test. Since this one single case is the case of local linear estimation with the Anderson-Darling distance measure, it makes sense to conclude that there is no evidence against positive dependence.

	KS	CvM	AD
E	.255	.084	.064
LL	.115	.074	.049
MR	.099	.086	.087
LLS	.144	.081	.084
MRS	.143	.081	.061
B	.549	.113	.088
M	.247	.166	.217

Table 2.14: *Approximated P-values for the data set of log-returns of BEL20 vs. EUR/DOL.*

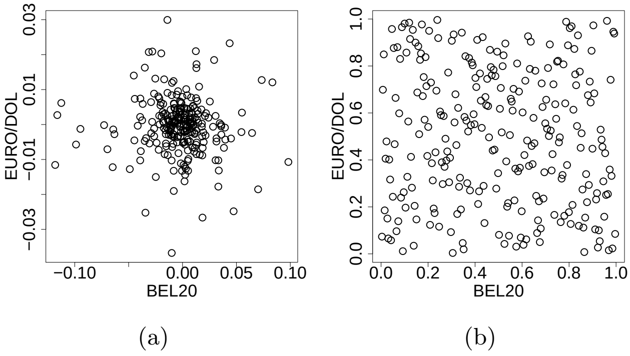


Figure 2.5: *Scatterplot of the data of log-returns of BEL20 vs. EUR/DOL (a) and plot of the corresponding pseudo-observations (b).*

The first two data sets were examples of clear positive quadrant dependence, with non rejected null hypothesis with very high p-values. In the first case it was a matter of strong positive quadrant dependence and large sample size, and in the second case of very strong positive quadrant dependence and a moderate sample size.

The third data set presented a less obvious case. In most of the tests we were close to the significance level of 5%. The Pearson's Chi-squared test do not reject that the data come from the Π copula.

2.5 Conclusions and further discussion

In this chapter we relied on recently developed copula estimators and on different functional distances to propose well performing testing procedures for testing the null hypothesis of positive quadrant dependence. We proved the consistency of the proposed tests and provided a simulation study to illustrate the finite sample performances on a diverse set of testing problems, including quite challenging problems. The testing procedures were illustrated on real data applications.

Other discrepancy measures than the three considered so far can be thought of. From our extensive study, we find it worth to report on the following modifications of the Cramér-von Mises (CvM) and the Anderson-

Darling (AD) distance measures:

$$S_n^{\text{CvM2}} = n \int_{\mathbb{I}^2} (uv - \hat{C}_n(u,v))_+^2 dudv \tag{2.7}$$

$$S_n^{\text{AD2}} = n \int_{\mathbb{I}^2} \frac{(uv - \hat{C}_n(u,v))_+^2}{uv(1-u)(1-v)} dudv. \tag{2.8}$$

A selection of simulation results for these alternative distances are provided in Tables 2.15—2.18, which present parts of Tables 2.2, 2.5, 2.9 and 2.10 extended with the simulation results for the distance measures in (2.7) and (2.8). Note the considerable improved power for the tests based on these distances for the classical copulas, as well as the “switch” in performance between the bootstrap and multiplier based tests. The multiplier method gives for CvM2 and AD2 based test statictics comparable powers to the empirical method for classical copulas, but still has far lower powers for the mixture copulas. The main findings from all simulations can be summarized as follows:

- The use of copula kernel type estimators increases the power of the testing procedures in particular for Kolmogorov-Smirnov based tests. For the other test statistics they only lead to slightly higher powers for “classical” copulas, but can possibly lead to a power loss for “mixture” alternatives.

- When using the empirical copula estimate (E), the Cramér-von Mises (CvM) and the Anderson-Darling (AD) based statistics give in all the cases higher power than the Kolmogorov-Smirnov (KS) based tests in all considered situations. Note that, the modified distance versions of these tests lead to considerably higher powers.

- The proposed resampling from an independence copula method seems to work comparatively with a bootstrap (B) method for classical families of copulas when using CvM and AD and the empirical copula. For CVM2 and AD2 our method works slightly better. For mixture copulas our method gives significantly higher power than bootstrap or multiplier methods for CvM and AD. The method of resampling from the independent copula increases the power of KS based tests.

In the simulation study, we obtained much better results for the “classical” copula families than for the more challenging “mixtures” of copulas. An issue that we did not discuss is the choice of the smoothing parameter h_n for the kernel estimators. We applied the data-driven method proposed by Omelka et al. (2009), where the bandwidth is selected as a minimizer of the integrated asymptotic mean squared error of the copula estimator using the Frank copula as a reference copula. This reference copula is fitted via the empirical Kendall's tau from the data. In the case of non-PQD examples which are oscillating around the independence copula, the empirical Kendall's tau is close to zero, and as a consequence the chosen data-driven bandwidth is very likely large and not very appropriate. This problem deserves further research and other methods of bandwidth selection, tailored also for testing purposes, should be developed.

The testing procedures described in this chapter use Monte Carlo simulation with the independence case as a reference case to obtain the critical values. Overall these tests outperform tests based on bootstrap or multiplier techniques.

Another possible approach would be to resample from a non-parametric copula estimate that satisfies the null hypothesis. This could lead to increased powers. This approach is the subject of Chapter 3.

	Frank				FGM				Gumbel			
	CvM	CvM2	AD	AD2	CvM	CvM2	AD	AD2	CvM	CvM2	AD	AD2
					$\tau = -0.11$							
E	.729	.741	.719	.739	.726	.745	.749	.757	.710	.705	.773	.755
LL	.742	.747	.543	.727	.762	.761	.536	.747	.765	.757	.791	.791
MR	.743	.750	.745	.751	.741	.751	.742	.753	.698	.693	.693	.693
LLS	.747	.752	.696	.747	.745	.753	.719	.765	.736	.735	.803	.776
MRS	.745	.755	.738	.748	.745	.752	.765	.764	.728	.731	.783	.764
B	.752	.694	.746	.663	.751	.703	.765	.676	.715	.638	.780	.674
M	.573	.740	.469	.737	.569	.746	.460	.749	.540	.703	.522	.744

Table 2.15: *Simulation results on classical copulas: extension of part of Table 2.2.*

	CvM	CvM2	AD	AD2	CvM	CvM2	AD	AD2	CvM	CvM2	AD	AD2
	θ =1.47				θ=9.5				θ =17.5			
E	.050	.051	.054	.049	.389	.315	.617	.512	.777	.656	.926	.855
LL	.044	.048	.064	.052	.261	.218	.802	.518	.520	.398	.969	.880
MR	.050	.050	.049	.048	.108	.098	.106	.103	.153	.111	.163	.133
LLS	.050	.049	.052	.051	.276	.266	.571	.468	.578	.513	.880	.814
MRS	.050	.051	.048	.050	.269	.264	.527	.453	.561	.507	.853	.789
B	.055	.037	.055	.031	.291	.171	.496	.250	.639	.378	.872	.541
M	.016	.052	.007	.048	.143	.227	.194	.318	.455	.470	.594	.643

Table 2.16: *Simulation results on mixtures on Frank copulas: extension of part of Table 2.5.*

	$\theta = -0.5$				$\theta = -0.2$				$\theta = 0.2$			
	CvM	CvM2	AD	AD2	CvM	CvM2	AD	AD2	CvM	CvM2	AD	AD2
E	.746	.595	.805	.647	.081	.076	.092	.084	.041	.039	.040	.041
LL	.635	.574	.836	.667	.082	.078	.130	.093	.036	.037	.076	.045
MR	.564	.531	.555	.535	.071	.071	.072	.073	.039	.040	.038	.039
LLS	.644	.584	.770	.656	.075	.074	.101	.088	.040	.037	.051	.041
MRS	.638	.582	.716	.638	.075	.074	.092	.087	.039	.037	.044	.040
B	.673	.505	.755	.495	.084	.056	.097	.053	.043	.026	.046	.025
M	.522	.557	.508	.570	.038	.071	.020	.073	.014	.037	.008	.039

Table 2.17: *Simulation results on Mardia copulas: extension of part of Table 2.9.*

	$\gamma = 8.056$				$\gamma = 11.944$				$\gamma = 17.560$			
	CvM	CvM2	AD	AD2	CvM	CvM2	AD	AD2	CvM	CvM2	AD	AD2
E	.542	.351	.710	.449	.897	.661	.965	.760	.999	.930	1	.972
LL	.387	.324	.791	.504	.664	.568	.963	.762	.913	.824	.996	.960
MR	.285	.252	.281	.263	.491	.429	.487	.445	.755	.683	.751	.707
LLS	.396	.335	.657	.449	.712	.601	.903	.735	.946	.870	.992	.947
MRS	.394	.333	.565	.430	.702	.596	.849	.721	.945	.871	.989	.942
B	.449	.253	.632	.286	.796	.496	.913	.550	.986	.835	.996	.868
M	.287	.306	.293	.353	.678	.566	.677	.618	.964	.885	.965	.904

Table 2.18: *Simulation results on mixtures of Mardia copulas: extension of part of Table 2.8.*

2.6 Proof

Proof of Theorem 2.1. For part (i) we need to prove that, if H_0 is true, then $\lim_n \mathbb{P}(\text{reject } H_0) \leq \alpha$.

Let $\epsilon > 0$ be fixed and consider $\delta > 0$ and $\eta > 0$, to be specified later on. Denote

$$\begin{aligned}
I_0 &= \{(u,v): \; C(u,v) = uv\} \\
I_1^\delta &= \{(u,v): \; uv < C(u,v) \leq uv + \delta\} \\
I_2^\delta &= \{(u,v): \; uv + \delta < C(u,v)\}.
\end{aligned}$$

Further, let S_n^A denote the test statistic restricted to a subset $A \subset \mathbb{I}^2 = [0,1] \times [0,1]$, e.g., $S_n^{A,\mathrm{KS}} = \sqrt{n} \sup_{(u,v)\in A}(uv - \hat{C}_n(u,v))_+$.

We further distinguish between the cases of the Cramér-von-Mises and the Anderson-Darling distance measures on the one hand, and the Kolmogorov-Smirnov distance measure on the other hand.

- For the Cramér-von Mises and the Anderson-Darling distance measures we proceed as follows. For brevity of presentation we only give details for the Cramér-von Mises distance, the one for the Anderson-Darling distance being along the same lines.

  First note that

$$\begin{aligned}
\mathbb{P}\left(S_n > c^{\Pi}_{\alpha,n}\right) &= \mathbb{P}\left(S_n^{I_0} + S_n^{I_1^\delta} + S_n^{I_2^\delta} > c^{\Pi}_{\alpha,n} - \eta + \eta\right) \leq \\
&\leq \mathbb{P}\left(S_n^{I_0} > c^{\Pi}_{\alpha,n} - \eta\right) + \mathbb{P}\left(S_n^{I_1^\delta} > \eta\right) + \mathbb{P}\left(S_n^{I_2^\delta} > 0\right). \qquad (2.9)
\end{aligned}$$

  Recall that G_C denotes the limiting Gaussian process of $\sqrt{n}(\hat{C}_n - C)$ and put $c^{\Pi}_{\alpha} := \lim_n c^{\Pi}_{\alpha,n}$. On I_0, $C(u,v) = \Pi(u,v)$ and from (2.3) and (2.4) it can be deduced that the covariance functions of the processes G_C and G_Π coincide on the interior of I_0 denoted by $\mathrm{int}(I_0)$ (because of assumption **(C1)**), which imply that the processes of G_C and G_Π for $(u,v) \in \mathrm{int}(I_0)$ have the same distribution. The assumption on ∂I_0 (see **(C4)**) ensures that zero probability is given to this set. These considerations justify the passage from integrals involving G_C and dC to integrals involving G_Π and $d\Pi$ in the bound of the first component in (2.9) below. Therefore, with the help of the weak convergence of the process $\sqrt{n}(\hat{C}_n - C)$ and assumption **(C4)** one obtains, for all sufficiently large n, the bound

$$
\begin{aligned}
&\mathbb{P}\left(S_n^{I_0} > c_{\alpha,n}^{\Pi} - \eta\right) \\
&= \mathbb{P}\left(n \iint_{I_0} \left(C(u,v) - \hat{C}_n(u,v)\right)_+^2 d\hat{C}_n(u,v) > c_{\alpha,n}^{\Pi} - \eta\right) \\
&\leq \mathbb{P}\left(\iint_{I_0} (-G_C(u,v))_+^2\, dC(u,v) > c_\alpha^{\Pi} - 2\eta\right) + \epsilon \\
&= \mathbb{P}\left(\iint_{I_0} (-G_\Pi(u,v))_+^2\, d\Pi(u,v) > c_\alpha^{\Pi} - 2\eta\right) + \epsilon \\
&\leq \mathbb{P}\left(\iint_{\mathbb{I}^2} (-G_\Pi(u,v))_+^2\, d\Pi(u,v) > c_\alpha^{\Pi}\right) + 2\epsilon \\
&= \alpha + 2\epsilon.
\end{aligned}
$$

The second component in (2.9) can be bounded for all n sufficiently large and for small enough δ

$$
\begin{aligned}
\mathbb{P}\left(S_n^{I_1^\delta} > \eta\right) &\leq \mathbb{P}\left(n \iint_{I_1^\delta} \left(C(u,v) - \hat{C}_n(u,v)\right)_+^2 d\hat{C}_n(u,v) > \eta\right) \\
&\leq \mathbb{P}\left(\iint_{I_1^\delta} (-G_C(u,v))_+^2\, dC(u,v) > \eta\right) + \epsilon \leq 2\,\epsilon. \qquad (2.10)
\end{aligned}
$$

The last inequality in (2.10) may be justified as follows. As the limiting process G_C is centered and Gaussian, for each $\epsilon > 0$ there exists $K < \infty$ such that

$$
\mathbb{P}\left(\sup_{\mathbb{I}^2} |G_C(u,v)| > K\right) < \epsilon. \qquad (2.11)
$$

Thus with probability greater than $1 - \epsilon$ it holds that

$$
\iint_{I_1^\delta} (-G_C(u,v))_+^2\, dC(u,v) \leq K^2 \mu_C(I_1^\delta), \qquad (2.12)
$$

where μ_C is the measure associated with the copula C. Let $\{\delta_k,\ k \in \mathbb{N}\}$ be a sequence of positive numbers decreasing to zero. The definition of I_1^δ implies that $\bigcap_{k=1}^{\infty} I_1^{\delta_k} = \emptyset$. Now, the continuity of a measure (see e.g., Lemma 1.14 of Kallenberg 1997) yields

$$
\lim_{k\to\infty} \mu_C(I_1^{\delta_k}) = 0,
$$

which implies that $\mu_C(I_1^\delta)$ can be made arbitrary small, and hence the right-hand side of (2.12) can be made smaller than η by taking δ small enough. This together with (2.11) yields (2.10).

Finally, the third component in (2.9) can be bounded by ϵ for all n sufficiently large

$$\begin{aligned}
&\mathbb{P}\left(S_n^{I_2^\delta} > 0\right) \leq \mathbb{P}\left(\sup_{I_2^\delta}\left(uv - \hat{C}_n(u,v)\right) > 0\right) \\
&\leq \mathbb{P}\left(\sup_{I_2^\delta}\left(C(u,v) - \hat{C}_n(u,v)\right) > \delta\right) \\
&\leq \epsilon. \qquad (2.13)
\end{aligned}$$

- For the Kolmogorov-Smirnov distance measure we proceed as follows:

$$\mathbb{P}\left(S_n > c_{\alpha,n}^\Pi\right) \leq \mathbb{P}\left(S_n^{I_0 \cup I_1^\delta} > c_{\alpha,n}^\Pi\right) + \mathbb{P}\left(S_n^{I_2^\delta} > 0\right) . \qquad (2.14)$$

The first term on the right-hand side of (2.14) may be bounded as

$$\begin{aligned}
&\mathbb{P}\left(S_n^{I_0 \cup I_1^\delta} > c_{\alpha,n}^\Pi\right) = \mathbb{P}\left(\sqrt{n}\sup_{I_0 \cup I_1^\delta}\left(uv - \hat{C}_n(u,v)\right)_+ > c_{\alpha,n}^\Pi\right) \\
&\leq \mathbb{P}\left(\sqrt{n}\sup_{I_0 \cup I_1^\delta}\left(C(u,v) - \hat{C}_n(u,v)\right)_+ > c_{\alpha,n}^\Pi\right) \\
&\leq \mathbb{P}\left(\sup_{I_0 \cup I_1^\delta}\left(-G_C(u,v)\right)_+ > c_\alpha^\Pi - \eta\right) + \epsilon \qquad (2.15)
\end{aligned}$$

Let $\{\delta_k, k \in \mathbb{N}\}$ be again a decreasing sequence of positive numbers going to zero and for $k \in \mathbb{N}$ put

$$A_k = \left[\sup_{I_0 \cup I_1^{\delta_k}}\left(-G_C(u,v)\right)_+ > c_\alpha^\Pi - \eta\right].$$

Note that $A_k \supset A_{k+1}$ and by the almost sure continuity of the paths of the process G_C (see e.g., Addendum 1.5.8 of van der Vaart & Wellner 1996) we have

$$\bigcap_{k=1}^\infty A_k = \left[\sup_{I_0}\left(-G_C(u,v)\right)_+ \geq c_\alpha^\Pi - \eta\right] \cup N, \quad \text{where } \mathbb{P}(N) = 0.$$

Once more using the continuity of the probability measure from above we get

$$\lim_{k\to\infty} \mathbb{P}\left(\sup_{I_0 \cup I_1^{\delta_k}} (-G_C(u,v))_+ > c_\alpha^\Pi - \eta \right) = \mathbb{P}\left(\sup_{I_0} (-G_C(u,v))_+ \geq c_\alpha^\Pi - \eta \right).$$

Thus the probability on the right-hand side of (2.15) may be, for sufficiently small δ and η, bounded as

$$\begin{aligned}
&\mathbb{P}\left(\sup_{I_0 \cup I_1^{\delta}} (-G_C(u,v))_+ > c_\alpha^\Pi - \eta \right)\\
&\leq \mathbb{P}\left(\sup_{I_0} (-G_C(u,v))_+ \geq c_\alpha^\Pi - \eta \right) + \epsilon\\
&\leq \mathbb{P}\left(\sup_{I_0} (-G_C(u,v))_+ \geq c_\alpha^\Pi \right) + 2\epsilon\\
&= \mathbb{P}\left(\sup_{I_0} (-G_\Pi(u,v))_+ \geq c_\alpha^\Pi \right) + 2\epsilon\\
&\leq \mathbb{P}\left(\sup_{\mathbb{I}^2} (-G_\Pi(u,v))_+ \geq c_\alpha^\Pi \right) + 2\epsilon = \alpha + 2\epsilon, \qquad (2.16)
\end{aligned}$$

which together with (2.15) gives for all sufficiently large n

$$\mathbb{P}\left(S_n^{I_0 \cup I_1^\delta} > c_{\alpha,n}^\Pi \right) \leq \alpha + 3\epsilon.$$

The first equality in (2.16) follows from the fact that according to the remark after Theorem 2.1, the limiting processes G_C and G_Π are zero on the boundary of $[0,1]^2$, and therefore it suffices to look at the supremum over the set $I_0 \cap \operatorname{int}([0,1]^2) = I_0 \cap (0,1)^2$. Below we show that for all points of this set it holds that

$$c_u(u,v) = v = \Pi_u(u,v) \qquad \text{and} \qquad c_v(u,v) = u = \Pi_v(u,v)\ , \quad (2.17)$$

which then implies that the limiting processes G_C and G_Π have the same distribution on $I_0 \cap (0,1)^2$. For proving statement (2.17) suppose there is a point $(u_0, v_0) \in I_0 \cap (0,1)^2$, such that for instance $c_u(u_0, v_0) <$

v_0. Put $\varepsilon = v_0 - c_u(u_0, v_0)$. Then by Taylor expansion for a sufficiently small $\Delta > 0$, we have that

$$\begin{aligned}&C(u_0 + \Delta, v_0) = C(u_0, v_0) + \Delta\, c_u(u_0, v_0) + o(\Delta)\\ &= u_0\, v_0 + \Delta\, c_u(u_0, v_0) + o(\Delta) < u_0\, v_0 + \Delta\,(v_0 - \tfrac{\varepsilon}{2})\\ &< (u_0 + \Delta)\, v_0 = \Pi(u_0 + \Delta, v_0),\end{aligned}$$

which contradicts the null hypothesis of PQD. A similar argument can be given if $C_u(u_0, v_0) > v_0$ or $C_v(u_0, v_0) \neq u_0$. This completes the proof of (2.17) and the bound for the first component in (2.14).

The second component in (2.14) can be bounded by ϵ for all n sufficiently large in an analogous manner as in (2.13).

For part (ii) of Theorem 2.1 we need to prove that, if H_0 is false, then $\lim\limits_{n\to\infty} \mathbb{P}(\text{reject } H_0) = 1$.
Since it holds that

$$\exists (u, v) \in \mathbb{I}^2 : \quad C(u, v) < uv,$$

and because of the continuity of the distances, the test statistic S_n converges to infinity in probability, so

$$\mathbb{P}(S_n^C > c_{\alpha,n}^{\Pi}) \longrightarrow 1.$$

$\square$

Chapter 3

Constrained copula estimation for positive quadrant dependence testing

3.1 Introduction

This chapter is based on Gijbels and Sznajder (2011b) and develops further tests for positive quadrant dependence.

In the previous chapter we tested for PQD in the data focusing on the characterization (1.9) and testing the null hypothesis

$$H_0 : \ \forall u, v \in [0,1] \qquad C(u,v) \geq \Pi(u,v)$$

versus

$$H_1 : \ \exists u, v \in [0,1] \qquad C(u,v) < \Pi(u,v).$$

The test statistics were built upon a functional distance between a copula estimator and the independence copula function Π. We compared two approaches to obtain the critical values of such tests. The first approach relied on approximations of the asymptotic distribution of the considered test statistic (Scaillet (2005)). The second approach, introduced and studied by Gijbels et al. (2010), used an approximated finite-sample distribution of the test statistic under the reference copula distribution Π. This one turned out to lead to better performances for the considered tests.

Both existing approaches however do not (fully) exploit the fact that under the null hypothesis the copula is a PQD copula and satisfies constraint (1.9). In this chapter we introduce a different finite-sample approach to remedy for this general drawback, and compute the critical values of the tests based on a constrained copula estimation. Specifically, we approximate the distribution of a test statistic for a copula distribution from the null set, which also resembles the copula shape given by the data. We propose two different ways to do this. The full exploitation of the null hypothesis leads to improved testing procedures. The key issue is to resample from a constrained copula estimation. This issue appears to be a real challenge since: (i) most non-parametric copula estimators are no real copulas themselves (i.e., are not satisfying all requirements for a copula function) and the same holds for constrained copula estimators; (ii) it is not clear how to resample from a constrained copula estimator. In this chapter we deal with these two challenges.

This chapter is organized as follows. In Section 3.2 we discuss existing methods for the PQD testing and motivate the introduction of the new methods. In Section 3.3 we describe the new testing procedures and their properties. Section 3.4 contains the results of a simulation study that investigates the finite-sample power and size performances of the proposed new methods and compares them with the performances of existing testing procedures. In Section 3.5 the PQD testing procedures are applied to the Danish fire insurance data. Finally, Section 3.6 gives some conclusions and further discussions.

3.2 Testing for PQD

Recall that Scaillet (2005) investigates the asymptotic distribution of the test statistic, built upon the Kolmogorov-Smirnov distance and relies on the weak convergence result of the empirical copula process $\sqrt{n}\big(C_n - C\big)$ (see Fermanian et al. (2004)), which is also extensively used in goodness-of-fit tests for copulas (see Genest et al. (2009b) and Kojadinovic and Yan (2011)). This process converges to a centered Gaussian process whose covariance structure not only depends on the values of the copula itself, but also on its partial derivatives. To approximate this limiting distribution the so called bootstrap and multiplier methods are used. Eventually, the obtained resampled process is used to approximate the limiting distribution of the test statistic and to compute the critical value for the test.

A different approach was proposed in Chapter 2 and was based on a reference copula. There, instead of approximating the asymptotic distribution of a test statistic, the critical values are obtained by approximating the finite-sample distribution of the test statistic under the null hypothesis from a reference copula distribution. A natural candidate for such a reference distribution in the PQD testing problem is the independence copula Π, as it is the boundary element in the null set. The procedure was to randomly draw a multitude N of independent samples of size n from the independence copula $\left\{(U_i^{*(k)}, V_i^{*(k)})_{i=1,\ldots,n}\right\}_{k=1}^N$, where $(U_i^{*(k)}, V_i^{*(k)}) \sim_{iid} \Pi$ and used them to calculate the corresponding test statistic values $\{S_{n,\Pi}^{*(k)}\}_{k=1}^N$, which then lead to an approximation of the distribution of the test statistic under Π and to the critical values of the test, namely

$$c_{\alpha,\Pi}^N = F_{S_{n,\Pi}^*,N}^{-1}(1-\alpha), \tag{3.1}$$

where $F_{S_{n,\Pi}^*,N}^{-1}$ is the empirical quantile function built from $\{S_{n,\Pi}^{*(k)}\}_{k=1}^N$. We shall drop the superscript N from the critical value notation, as the approximation error can be reduced by increasing N. Obviously, this approach is not fully exploiting the null hypothesis setting, since it makes reference to one particular element (namely Π) in H_0.

Both existing approaches have a drawback. In the first approach, the asymptotic test statistic distribution is not approximated under the null hypothesis. This leads to a rather low power, as was noticed in the simulation study in Section 2.3. In the second approach, although the test statistic distribution is approximated under the null hypothesis, it does not take the data into account in the approximation process, since the same reference Π is always used. The methods proposed in Section 3.3 try to deal with these drawbacks, by obtaining a test statistic distribution under the null hypothesis, which also takes into account the fact that a copula can violate the PQD condition only on a subset of the unit square. This requires a constrained copula estimator and a resampling procedure. In the next section we explain how we deal with these issues.

3.3 Constrained copula estimation and PQD testing

The PQD testing procedure proposed here consists of two parts: a constrained copula estimation and a resampling process.

A PQD-constrained copula estimator should yield a consistent estimator under the null hypothesis and a PQD copula under the alternative. In the literature, Ebrahimi (1993) described a procedure of reweighting the empirical copula estimator to create a PQD-restricted estimator on a grid of points built on the observation points. This estimator requires however solving a constrained high-dimensional minimization problem (with n^2 parameters), which even for a moderate sample size is rather unfeasible, and therefore will not be further discussed here.

A general resampling process from a copula requires finding a conditional distribution of one variable given another, as expressed in Theorem 1.2. As the partial derivative of any copula function exists almost everywhere, the resampling can be done according to Algorithm 1.1. This very general method allows for a similar resampling procedure as in the Π-reference case (that uses (3.1)) to obtain the critical values, yet where S_n^* comes not from the Π distribution, but from a PQD copula distribution shaped by the original sample.

In the following parts we propose two such finite-sample procedures: a reference to a constrained non-parametric estimation and a reference to a parametric copula family with a PQD-constrained estimation within that family.

3.3.1 PQD-constrained non-parametric estimation

If we have a smooth PQD-constrained non-parametric copula estimator, then we can directly resample from it by using Algorithm 1.1. However, if we do not have such an estimator, then we can first get any PQD-constrained copula estimator and possibly smooth it to obtain the partial derivative.

The idea is to obtain a PQD copula estimator which will converge to a PQD copula, which is closest, in some sense, to the true underlying copula. We obtain this by defining an operator from the set of copula estimators into the subset of PQD copula estimators, and which is invariant on this subset (i.e., if the copula estimator is PQD then the operator does not alter the estimator). We start by motivating such an operator on the level of theoretical copula functions.

Theoretical motivation

Let us consider the following operator on the set of copulas

$$C^+ = \max(C, \Pi), \tag{3.2}$$

where the maximum is taken pointwise. The resulting function is a bivariate function on the unit square, but which is not necessarily a copula. In Proposition 3.1 we present some special cases when C^+ is indeed a copula function.

Proposition 3.1. *Let C be a copula function.*

1. *If C is a PQD copula, then C^+ is also a PQD copula.*
2. *If C^+ is a copula, then it is also a PQD copula.*
3. *Denote by $I_C = \{(u,v) \in [0,1]^2 : \; C(u,v) \geq \Pi(u,v)\}$ and let ∂I_C be the boundary of this set, and μ_{C^+} the measure on the unit square generated by C^+ (originating from the C^+-volumes of rectangles). Then we have, that if $\mu_{C^+}(\partial I_C) \geq 0$, then C^+ is a copula.*

Proof. The first two points in Proposition 3.1 follow immediately from definition (3.2).
The third point comes from the fact that, if C^+ is not a copula, then it also does not define a proper probability measure on the unit square, specifically μ_{C^+} is a signed measure. Furthermore, the only negative mass can be contained in ∂I_C. So excluding this to happen ensures that C^+ is indeed a copula. □

The first two points of Proposition 3.1 provide in fact crucial features for the constrained estimation. Firstly, if C is already a PQD copula, then we want $C^+ \equiv C$, which is the case. Secondly, if C^+ is a copula, then it is the closest copula to C among the PQD copulas, i.e.,

$$\forall \bar{C} \in \mathcal{C}^+ \quad ||C - C^+||_{L_p([0,1]^2)} \leq ||C - \bar{C}||_{L_p([0,1]^2)} \;, \tag{3.3}$$

where $\mathcal{C}^+$ is the set of all PQD copulas, $p \geq 1$, and, where $\|\cdot\|_{L_p(A)}$ denotes the L_p norm on the set A. Indeed, as

$$\iint_{I_C} |C(u,v) - C^+(u,v)|^p du dv = 0 \leq \iint_{I_C} |C(u,v) - \bar{C}(u,v)|^p du dv$$

and

$$\begin{aligned}
&\iint_{[0,1]^2 \setminus I_C} |C(u,v) - C^+(u,v)|^p du dv \\
&= \iint_{[0,1]^2 \setminus I_C} |C(u,v) - \Pi(u,v)|^p du dv \\
&\leq \iint_{[0,1]^2 \setminus I_C} |C(u,v) - \bar{C}(u,v)|^p du dv \;,
\end{aligned}$$

statement (3.3) follows.

The third point in Proposition 3.1 describes a broad class of copulas for which the $\cdot^+$ operator gives a valid copula. One special example of copulas such that $\mu_{C^+}(\partial I_C) = 0$ is the collection of copulas for which ∂I_C consist of intervals orthogonal to the axes and $\mu_C(\partial I_C) = 0$.

For the remaining copulas, for which C^+ is not a copula, we extend the $\cdot^+$ operator to an operator $\cdot^{++}$ and conjecture for this that it gives a copula that is "almost" PQD. The $\cdot^{++}$ operator will be based on the increasing rearrangement technique, described in Definition 1.7, adapted to our case.

Let us consider the partial derivative of $C^+(u, v)$ with respect to u

$$c_u^+(v) = \frac{\partial C^+(u, v)}{\partial u}. \tag{3.4}$$

Then let us apply the monotone rearrangement technique to it

$$\widetilde{c}_u^+(v) = (\xi \circ \xi)\big(c_u^+(\cdot)\big)(v). \tag{3.5}$$

Then

$$\widetilde{C}^+(u, v) = \int_0^u \widetilde{c}_x^+(v)dx. \tag{3.6}$$

is a bivariate distribution function on the unit square, such that

$$\widetilde{C}^+ \sim (U, \widetilde{V}), \tag{3.7}$$

where $U \sim U[0, 1]$ and $\widetilde{V} \sim \widetilde{G}$, where

$$\widetilde{G}(v) = \int_0^1 \widetilde{c}_x^+(v)dx = \widetilde{C}^+(1, v). \tag{3.8}$$

Thus, then there exists a unique copula $C_{U,\widetilde{V}}$ (denoted by C^{++})

$$C^{++}(u, v) = \widetilde{C}^+\big(u, \widetilde{G}^{-1}(v)\big). \tag{3.9}$$

Note that if C^+ is a copula, then $c_u^+(v)$ is just its conditional distribution function, like in Theorem 1.2. Thus it is a non-decreasing function and according to Proposition 1.7 (c) $\widetilde{c}_u^+ \equiv c_u^+$ for all u and as a consequence $C^{++} \equiv C^+$.

Further we can see that $\widetilde{C}^+$, defined in (3.6), is a proper bivariate distribution function even if C^+ is not a copula (and hence c_u^+ is not necessarily non-decreasing). Indeed, the transformation ξ is continuous on $L_\infty([0, 1])$.

Moreover, it is unaffected by changing function values on sets with zero Lebesgue measure. Hence, $\widetilde{c}_u^+(v)$ is well defined everywhere. For notational correctness we can define $c_u^+(v)$ to be equal to $c_u(v)$, whenever the former does not exists from the definition (3.4).

Moreover, $\widetilde{C}^+$ lies in some sense closest to C^+ according to Proposition 1.7 (d). More precisely, let H^* denote the continuous bivariate distribution function with the first marginal being uniformly distributed and with no mass on the boundaries of the unit square which is "closest" to C^+ function. Then, by applying rearangement technique, $\widetilde{C}^+$ is "closer" to H^*, than H^* is to C^+.

Proposition 3.2. *If H_* is any continuous bivariate distribution function on the unit square with the first marginal being uniformly distributed and with no mass at the boundaries of the unit square, then*

$$||\widetilde{C}^+ - H_*||_{L_\infty([0,1]^2)} \leq ||C^+ - H_*||_{L_\infty([0,1]^2)}.$$

Proof. (*Reductio ad absurdum*) Suppose that

$$||\widetilde{C}^+ - H_*||_{L_\infty([0,1]^2)} > ||C^+ - H_*||_{L_\infty([0,1]^2)}$$

and the supremums are realized in respective points (u_1, v_1) and (u_2, v_2). Thus,

$$\begin{aligned} |\widetilde{C}^+(u_1, v_1) - H_*(u_1, v_1)| &> |C^+(u_2, v_2) - H_*(u_2, v_2)| \\ &\geq |C^+(u_1, v_1) - H_*(u_1, v_1)|. \end{aligned}$$

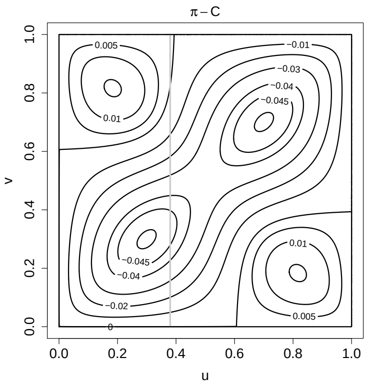
Let us assume, for example, that $\widetilde{C}^+(u_1, v_1) > H_*(u_1, v_1)$ and $H_*(u_1, v_1) \leq C^+(u_1, v_1)$ and define functions $f(u) = \widetilde{C}^+(u, v_1) - H_*(u, v_1)$ and $g(u) = C^+(u, v_1) - H_*(u, v_1)$. Then it follows that $f(u_1) > g(u_1) \geq 0$ and also $f(0) = g(0) = 0$, which implies that there exist a value u_3 such that

$$f'(u_3) > g'(u_3) > 0.$$

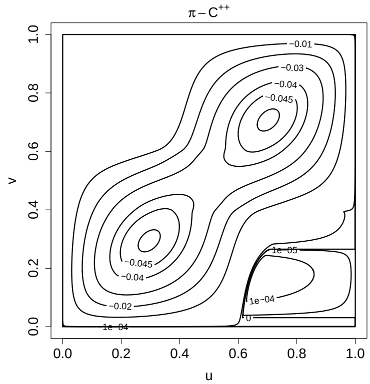
This contradicts Proposition 1.7 (d). □

Note that this proposition holds also for H being a copula function. Thus, if we denote by C^* the copula function which is the closest to C^+, then

$$||\widetilde{C}^+ - C^*||_{L_\infty([0,1]^2)} \leq ||C^+ - C^*||_{L_\infty([0,1]^2)}.$$



(a)



(b)

Figure 3.1: *Contour plots. (a): for* $\Pi - C_{mF}$*; (b): for* $\Pi - C_{mF}^{++}$.

We provide an example to explain the procedure and to further support the reasoning behind the conjecture. Let us look at a mixture of two Frank copulas (see Section 1.2.2 for the definition of a Frank copula). Specifically,

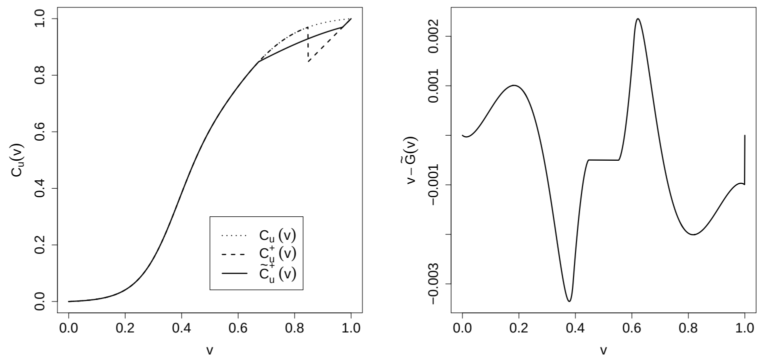
$$C_{\mathrm{mF}} = 0.6C_{\mathrm{Frank}(10)} + 0.4C_{\mathrm{Frank}(-15)}.$$

In Figure 3.1(a) we depict contour plots of the difference between the Π copula and C_{mF}. The C_{mF} copula is clearly not PQD. In Figure 3.1(b) we present the differences between the Π copula and the transformed copula C_{mF}^{++}. Note that everywhere we have negative values (except at the lower-right corner where the differences are of negligible orders such as 10^{-4} and 10^{-5}). The transformed copula C_{mF}^{++} is much closer to being PQD, specifically the positive hills in the upper-left and lower-right corners have diminished almost completely.

In Figure 3.1 we also indicate with a grey vertical line a u-section for which we depict the partial derivatives $c_{\mathrm{mF},u}(v)$, $c_{\mathrm{mF},u}^{+}(v)$ and $\widetilde{c}_{\mathrm{mF},u}^{+}(v)$ in Figure 3.2(a). This figure clearly illustrates the monotone rearrangement technique applied to our context. Figure 3.2(b) contains an approximation of $v - \widetilde{G}_{\mathrm{mF}}(v)$ from (3.8), which suggests that $\widetilde{V}_{\mathrm{mF}}$ from (3.7) has a distribution close to the uniform one.

Estimator formulation

If $\hat{C}$ is a smooth consistent copula estimator, then $\hat{C}^{++}$ is a PQD-constrained non-parametric copula estimator.



(a) (b)

Figure 3.2: *(a). Partial derivatives $c_{mF,u}(v)$ (dotted curve), $c^{+}_{mF,u}(v)$ (dashed curve) and $\widetilde{c}^{+}_{mF,u}(v)$ (solid curve); (b). The difference $v - \widetilde{G}_{mF}(v)$.*

For a non-smooth copula estimator, e.g., the empirical copula estimator, the estimation procedure requires an additional step of pre-smoothing, so that (3.4) is well defined. For simplicity of notation we will still denote such an operator as $^{++}$, e.g., C_n^{++} will denote a PQD-constrained non-parametric copula estimator based on the empirical copula estimator (but with pre-smoothing step).

For the pre-smoothing step, we propose an approach by means of local polynomial techniques. The idea is to approximate the empirical version of the surface (3.2) on a grid and smooth it out, as described in Section 1.2.3. The number of grid points has to depend on the sample size, but for simplicity of presentation we drop this dependence on n from the notation.

Let us now apply (1.4) and put

$$c_{i,j} = \begin{cases} C_n^{+}(u_i, v_j) & i, j = 1, \ldots, m \;, \\ \Pi(u_i, v_j) & i \text{ or } j \in \{0, m+1\}. \end{cases}$$

In other words $c_{i,j}$ is equal to the empirical PQD surface (3.2) in the interior of the unit square and to the true values on the boundary of it, according to Proposition 1.2 (b).

The validity of the procedure is expressed in Theorem 3.1.

Theorem 3.1. *If* $m \to \infty$, $h_1, h_2 \to 0$, $m^2 h_1 h_2 \to \infty$, $m^2 h_1 h_2/\sqrt{n} \to 0$ *for* $n \to \infty$ *and some regulatory conditions on the kernel* k *are fulfilled such that the smallest eigenvalue of* $X'WX$ *is bounded from below for sufficiently large* m*, then* $\hat{c}_u^+(v)$ *is an asymptotically consistent estimator of* $c_u^+(v)$*, whenever the latter exists.*

Proof. (Sketch). The elements of the vector Y can be written as

$$c_{i,j} = C(u_i, v_j) + \varepsilon_n^{(i,j)},$$

where $\mathbb{E}\varepsilon_n^{(i,j)} = o\left(n^{-\frac{1}{2}}\right)$ and $\mathbb{E}\varepsilon_n^{(i,j)}\varepsilon_n^{(k,l)} = O\left(n^{-1}\right)$ uniformly in i, j, k, l. Now,

$$\mathbb{E}\hat{c}_u^+(v) = [0,1,0](X'WX)^{-1}X'W\underline{C} + [0,1,0](X'WX)^{-1}X'W\mathbb{E}\varepsilon_n$$

and

$$\mathbb{E}\big(\hat{c}_u^+(v)\big)^2 = [0,1,0](X'WX)^{-1}X'W\mathbb{E}\left(\varepsilon_n\varepsilon_n'\right)WX(X'WX)^{-1}[0,1,0]',$$

where

$$\underline{C} = [C(u_1, v_1), C(u_1, v_2), \ldots, C(u_m, v_m)]'$$

and

$$\varepsilon_n = [\varepsilon_n^{(1,1)}, \varepsilon_n^{(1,2)}, \ldots, \varepsilon_n^{(m,m)}]'.$$

As $(X'WX)^{-1}$ is bounded and $X'W\mathbb{E}\varepsilon_n$, and $X'W\mathbb{E}\left(\varepsilon_n\varepsilon_n'\right)WX$ converge to zero this concludes the proof. □

In our case the regulatory conditions could be that the kernel function k is bounded and has support on $[0;1]$, which is fulfilled, for example, by Epanechnikov kernel.

Once a smooth estimator of $c_u^+(v)$ is obtained we can apply transformations (3.5), (3.6) and (3.9) to obtain a complete non-parametric PQD-constrained copula estimator C_n^{++}.

Note that one can use any other consistent estimator, instead of taking the empirical copula estimator, and that one can use any smoothing technique that consistently approximates the partial derivative, instead of the local linear smoothing technique.

It is also important to note that, for the purpose of resampling from a constrained estimator, it suffices to finish the estimation procedure on transformation (3.5). In other words, $\widetilde{c}_u^+$ is enough to obtain pseudo-samples from C^{++}, as pseudo-observations are rank based, thus invariant to the probability transformation. However, $\widetilde{\hat{c}}_u^+$ (the monotonized $\hat{c}_u^+$) does not always have to be a valid cumulative distribution function on the unit interval, as opposed to its theoretical counterpart. Therefore, it is often necessary to apply a linear transformation to $\widetilde{\hat{c}}_u^+$

$$\widetilde{\widetilde{\hat{c}}}_u^+(v) = \frac{\widetilde{\hat{c}}_u^+(v) - \widetilde{\hat{c}}_u^+(0)}{\widetilde{\hat{c}}_u^+(1) - \widetilde{\hat{c}}_u^+(0)}.$$

Having an estimate of a partial derivative surface $\widetilde{\widetilde{\hat{c}}}_u^+(v)$, we can apply the standard approach in Algorithm (1.1) to generate bivariate samples which are approximately $\widetilde{C}^+$-distributed. Such generated samples are then used to approximate the distribution of the test statistics and consequently the p-values of the original sample test statistics as in (3.1), i.e.,

$$c_{\alpha,C^{++}} = F^{-1}_{S^*_{n,C^{++}}}(1-\alpha),$$

where $F^{-1}_{S^*_{n,C^{++}}}$ is the (empirical) quantile function of $S^*_{n,C^{++}}$, which is based on a resampled sample of size N from approximately C^{++}.

This constrained non-parametric approach can be seen as completely opposed to the Π-reference approach in terms of the resampling distribution. In the first case, the resampling distribution is a data-driven PQD-adjusted distribution, whereas in the second one, it is set to a fixed reference distribution. Since these are two extreme approaches, it is of interest to investigate an “intermediate” approach. Such an “intermediate” approach, based on a parametric copula family, is discussed in the next section.

3.3.2 PQD-constrained parametric estimation

Consider a parametric copula family $\mathcal{C}$ with a parameter space Θ, which might be of (a priori) interest to a researcher. If this family includes at least a nonempty subset of PQD copulas $\mathcal{C}^+ \subseteq \mathcal{C}$, we can fit a copula from that subset to the data and obtain a reference distribution for the resampling procedure and the critical values in a similar way as in (3.1).

For generic reasons, the subset $\mathcal{C}^+$ can be of an arbitrary content as well as the copula family $\mathcal{C}$. Yet it is reasonable to have a family $\mathcal{C}$ that includes a broad spectrum of PQD copulas ranging from the independence copula Π up to the Fréchet-Hoeffding upper bound copula M.

We consider here two examples of such copula families $\mathcal{C}$ coming from a class of one parametric Archimedean copula models introduced in Section 1.2.2, namely Frank and Clayton copula families. These examples of copula families have several convenient features. One of them is that there exists a bijection between the parameter θ and the Kendall's tau association measure, defined in Definition 1.10, which gives for these two copula families (see Nelsen (2006))

$$\tau_{\text{Frank}}(\theta) = 1 - \frac{4}{\theta}\left(1 - \frac{1}{\theta}\int_0^\theta \frac{t}{e^t - 1}dt\right)$$
$$\tau_{\text{Clayton}}(\theta) = \frac{\theta}{\theta + 2} .$$

This allows for a simple estimation method, based on the inversion of the empirical Kendall's tau τ_n defined in Definition 1.9.

$$\theta_n = \tau^{-1}(\tau_n). \tag{3.10}$$

Another important feature of these examples of copula families is that they are positively ordered with respect to the concordance order, i.e.,

$$\theta_1 \leq \theta_2 \implies C_{\theta_1}(u,v) \leq C_{\theta_2}(u,v) \;\; \forall u,v \in I . \tag{3.11}$$

As a consequence, it means that for both example families, $\mathcal{C}^+$ is characterized by the parameter subset $\Theta^+ = [0,\infty]$. Furthermore, the constrained copula estimation can take a simple form

$$\theta_n^+ = \max\left(0, \tau^{-1}(\tau_n)\right) . \tag{3.12}$$

Lastly, both families are comprehensive, i.e., they reach from the lower to the upper Fréchet-Hoeffding bounds passing through Π. Thus, they are able to model the span of the dependency structure. They are however, as every Archimedean copula, symmetric, which might be a limitation for certain data examples. This is, among others, explored in the simulation study in Section 3.4.

It is important to note that we give expressions (3.10) and (3.12) only as an example of a parametric PQD-constrained copula estimation. They originate, however, from a general idea that having any consistent estimator $\hat{\theta}_n$ of the parameter θ we can construct a constrained estimator from it by considering the following expression

$$\hat{\theta}_n^+ = \text{argmin}\left\{\theta^+ \in \Theta^+ : \; L(C_{\hat{\theta}_n}, C_{\theta^+})\right\},$$

where L is a functional distance, e.g., $L_2\big([0,1]^2\big)$. This approach is a special (parametric) case of the general concept of constrained projection estimators, see e.g., Fils-Villetard et al. (2008) for an application in constrained extreme-value copula estimation. Note also that in the non-parametric setting the $^+$ operator (3.2) is a projection operator, yet not onto the set of the PQD copulas. The $^{++}$ operator was defined as a step to correct for that.

Eventually, in Proposition 3.3 we state the theoretical requirement for the resampling procedure to work. Please note, that we present the result for a general situation of having a projection estimator. This requires the co-domain to be closed and convex, which is the case for PQD copulas.

Proposition 3.3. *Let $C_{X,Y} = C_{\theta_0}$ and $\theta_0 \in \Theta^+$. If $\hat{\theta}$ is a consistent estimator of θ and $^+$ is a projection operator (onto Θ^+), then $\frac{\partial C_{\hat{\theta}^+}(u,v)}{\partial u}$ is a consistent estimator of $\frac{\partial C_{\theta_0}(u,v)}{\partial u}$ if the latter is continuous in θ_0.*

Proof. The proof follows from the continuity of a projection operator. □

Note that if the first order partial derivative of C_θ is continuous on the whole unit square, then $\frac{\partial C_{\hat{\theta}^+}(u,v)}{\partial u}$ is a uniformly consistent estimator. This is the case for the Frank and Clayton copula families.

3.4 Simulation study

A Monte Carlo simulation study is conducted to investigate the finite-sample power performance of the proposed testing procedures. Several copula function models are considered. The sample size is 200, the number of samples used for approximating the power is 1000 and for each sample another 1000 samples are drawn to approximate the p-value. The significance level is set to 5%. In the non-parametric approach the number of grid points is obtained from $m = 14$ and we use bandwidths $h_1 = h_2 = 1.5/m$. The considered copula functions are partially taken from the simulation study in the previous chapter for comparison reasons. The simulation study is performed on the K.U.Leuven cluster vic3 with the usage of R software (R Development Core Team (2011)), in particular the *copula* R package (Yan (2007)).

We compare the performances of three testing procedures:

NON-PARAMETRIC: the proposed non-parametrically built testing procedure of Section 3.3.1;

PARAMETRIC: the proposed parametrically built testing procedure of Section 3.3.2 using two different parametric reference copula families (a Frank family and a Clayton family);

Π-REFERENCE: the Π-reference built testing procedure from Chapter 2.

Power results

The first two considered copula examples in the simulation study are a Frank copula with parameter -1 and a Clayton copula with parameter -0.2. In both cases such parameters give a value of Kendall's tau equal to -0.11, which reflects a mild negative dependence. Tables 3.1 and 3.2 summarize some important characteristics for the copula models appearing in this simulation study.

copula	parameter values	τ	$\lambda_2\{C \geq \Pi\}$
Frank	$\theta = -1$	-0.11	0
Clayton	$\theta = -0.2$	-0.11	0
mFrank I	$(\theta_1, \theta_2, \gamma) = (9.5, -9.5, 0.5)$	0	0.50
mFrank II	$(\theta_1, \theta_2, \gamma) = (10, -15, 0.6)$	0.11	0.69
asym. cop.	$(\theta_1, \theta_2) = (-0.12, 8.3)$	-0.02	0.38

Table 3.1: *Characteristics of copulas examples in the simulation study.*

Note that for these first two examples, referring to (3.11), the true underlying copula C always lies below the Π copula, so it represents an NQD structure. In the last column of Tables 3.1 and 3.2 we indicate the proportion of the surface of $[0,1]^2$ on which a copula C lies above the Π copula. We denote by $\lambda_2(A)$, with $A \subseteq [0,1]^2$, the Lebesgue measure of the set A.

In these two examples of NQD, because of the considered constraint in the copula estimation process, both the non-parametric and parametric approach asymptotically yield the Π copula as the resampling distribution (see also Figure 3.3 for the density of the parameter estimates for the 1000 simulations). We thus expect that the power performance will be comparable to that for the Π-reference approach. This is indeed observed in Table 3.3, that presents the proportion of times out of 1000 that the testing procedure rejects the null hypothesis of PQD.

Now let us discuss copula examples which are elements of the extended Mardia copula family, as defined in (1.2).

For the parameters $(\theta, \gamma) = (-0.5, 1)$ we again obtain theoretical Kendall's tau close to -0.11. This copula is neither PQD nor NQD. It lies above the Π copula only in 12.5% of the unit square region. For the CvM and AD distance-based test statistics, the non-parametric method works equally well when compared to the parametric models as well as to the Π-reference approach. However, for the KS distance-based test statistic, the non-parametric method works best. Moreover, the overall power is bigger when compared to the first two examples.

The next two parameter sets for (θ, γ), namely $(-0.5, 0.825)$ and $(-0.2, 8.056)$, yield theoretical Kendall's taus equal to -0.09 and -0.10 respectively. In Table 3.2 we summarize characteristics for the three Mardia copula examples considered, where the weights $\omega_{\cdot}$ were defined in (1.2).

copula	(θ, γ)	ω_W	ω_Π	ω_M	τ	$\lambda_2\{C \geq \Pi\}$
Mardia	$(-0.5, 1)$	0.187	0.750	0.063	-0.11	0.125
eMardia I	$(-0.5, 0.825)$	0.155	0.794	0.051	-0.09	0.125
eMardia II	$(-0.2, 8.056)$	0.193	0.678	0.129	-0.10	0.32

Table 3.2: *Characteristics of the Mardia mixture examples.*

The second set of parameters $(-0.5, 0.825)$ diminishes the impact of the boundary copulas in the Mardia copula and magnifies the impact of the Π copula (see Table 3.2). Thus, it makes the violation (from PQD) harder to detect, leading to a decrease of the power performance overall when compared to the previous Mardia example.

The third set of parameters $(-0.2, 8.056)$ yields a copula which lies in 32% of the unit square region above the Π copula. This set of parameters decreases the impact of the Π copula in the mixture (see Table 3.2). From Table 3.3 we see that for the KS- and CvM-based test statistics the power is further decreased. In this situation the non-parametric approach works better. Note that the largest gain in power is obtained when using the AD-based test statistic.

Now let us analyze copula examples which are mixtures of two members of the Frank family

$$C^{\mathrm{mFrank}}_{\theta_1,\theta_2,\gamma} = \gamma C^{\mathrm{Frank}}_{\theta_1} + (1-\gamma) C^{\mathrm{Frank}}_{\theta_2}.$$

For the first set of parameters $(\theta_1, \theta_2, \gamma) = (9.5, -9.5, 0.5)$ the theoretical Kendall's tau is zero and half of the copula function lies above the Π copula (see also Table 3.1). Thus, among the considered examples this one is the most difficult one to detect for the Π-reference approach. The non-parametric approach works similarly to the other non-NQD cases, except for the CvM distance, which gives clearly smaller power than in the other cases. The parametric approach yields even more power than the non-parametric one.

distance	non-parametric	Π-reference	parametric $\mathcal{C} =$ Frank	parametric $\mathcal{C} =$ Clayton
		$C = \text{Frank}(-1)$		
KS	**0.608**	0.599	0.599	0.599
CvM	0.695	0.702	**0.704**	**0.704**
AD	0.706	0.705	**0.709**	**0.709**
		$C = \text{Clayton}(-0.2)$		
KS	**0.573**	0.551	0.551	0.551
CvM	**0.723**	0.715	**0.723**	**0.723**
AD	**0.795**	0.775	0.783	0.783
		$C =$ Mardia		
KS	**0.636**	0.614	0.615	0.615
CvM	0.775	0.772	**0.779**	**0.779**
AD	**0.856**	0.849	0.855	0.853
		$C =$ extended Mardia I		
KS	**0.465**	0.460	0.460	0.461
CvM	0.615	0.613	**0.619**	**0.619**
AD	0.709	0.699	**0.710**	**0.710**
		$C =$ extended Mardia II		
KS	**0.440**	0.400	0.414	0.412
CvM	**0.596**	0.546	0.561	0.560
AD	**0.769**	0.709	0.754	0.750
		$C =$ mixture of Frank I		
KS	0.457	0.357	**0.580**	0.522
CvM	0.499	0.366	**0.653**	0.607
AD	0.739	0.600	0.896	**0.899**
		$C =$ mixture of Frank II		
KS	0.317	0.138	0.945	**0.859**
CvM	0.386	0.086	**0.986**	0.974
AD	0.642	0.330	0.994	**0.995**
		$C =$ asymmetric copula (3.13)		
KS	**0.386**	0.277	0.327	0.313
CvM	**0.320**	0.218	0.259	0.251
AD	**0.500**	0.351	0.426	0.433

Table 3.3: *Power performances of the three methods for a selection of copula models. The highest proportion per row is indicated in bold.*

The second parameter set $(10, -15, 0.6)$ yields a copula which lies in 69% of the unit square region above the Π copula and has the theoretical Kendall's tau close to 0.11. Thus, in comparison to the previous copula this copula is a bit more difficult to distinguish from a PQD copula. Therefore, on the one hand, we can observe a decrease in power in the non-parametric and Π-reference approaches, yet the loss is much lower in the non-parametric case. On the other hand, because the parametric models are built on the estimated Kendall's tau, the parametric resampling produces samples coming from copulas deep into the null hypothesis, so the power increases in this approach.

Finally, to illustrate that a good choice of the parametric model for resampling is influential while testing for PQD, we present an example of an asymmetric copula based on the construction of copulas with quadratic sections as defined in Proposition 1.6.

In this simulation study we use

$$\psi(v) = \theta_1(1 - v^2)\sin(v\theta_2) \tag{3.13}$$

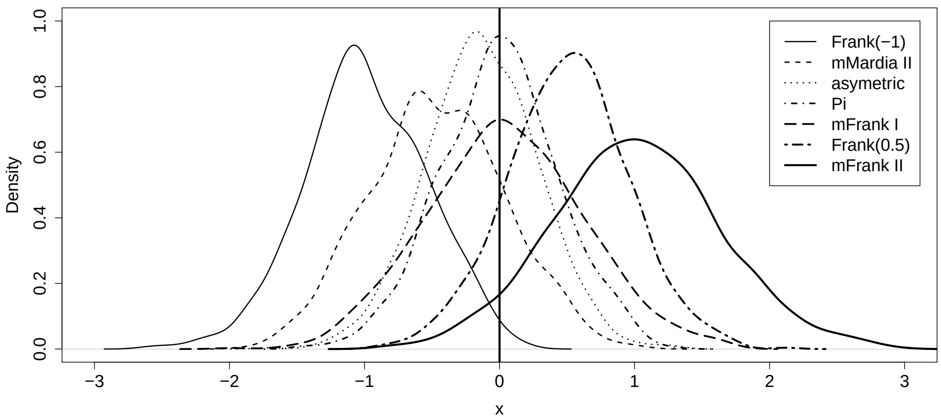
with the parameters $(\theta_1, \theta_2) = (-0.12, 8.3)$, which gives a copula with theoretical Kendall's tau around -0.02 and lying in almost 38% of the unit square above the independence copula (see also Table 3.1). In this example we clearly see, from Table 3.3, that the non-parametric method works best. The parametric models also work better than the Π-reference approach.

Overall, the proposed non-parametric and parametric built testing procedures behave in the same way as the Π-reference approach in case of the NQD copulas (the first two examples). The more complex the dependence structure the better the non-parametric approach works, as can be seen in the examples of the extended Mardia II and the asymmetric copulas. Furthermore, the parametric approach of Section 3.3.2 is always better than the Π-reference approach. As for the cross-distance analysis, the non-parametric approach improved the performance of the KS-based test statistic, yet overall the advisable choice is the AD-based test statistic.

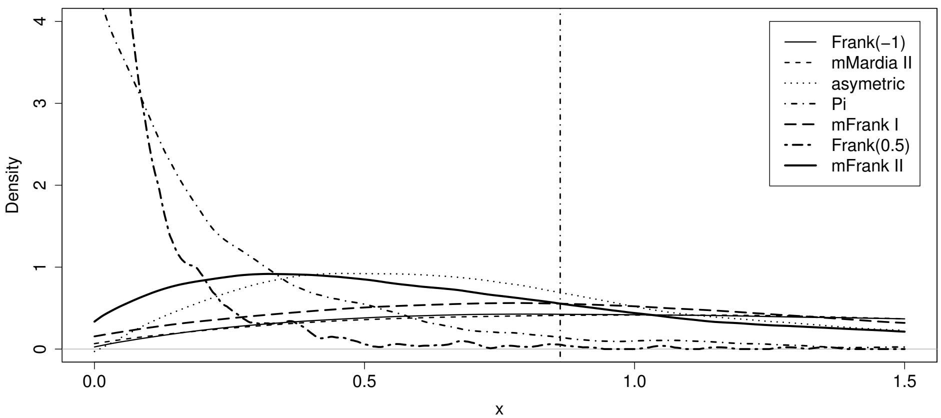
Finally, we provide Figure 3.3 to give more insight into the results of the parametric approach, seen from Table 3.3. In Figure 3.3(a), we depict approximated densities of the unconstrained parameter estimator $\hat{\theta}_n$ (based on the 1000 parameter estimates from the 1000 simulations) for a selection of the considered copula examples under the copula reference model $\mathcal{C}$ = Frank. Similar (not presented here) shape formations were captured under the copula model $\mathcal{C}$ = Clayton. We also do not include the examples with true copula C = Clayton, Mardia and extended Mardia I, as the estimated densities are similar to the case C = Frank(-1). We include instead two PQD copula model examples for comparison: Π and Frank(0.5). In Figure 3.3(b), we present the approximated densities of the AD-based test statistic for the corresponding copula examples. Additionally, we indicate the 0.95 quantile under the Π distribution.

From Figure 3.3(a) one can note that the density of the estimated parameter values for the mixture of Frank I lies symmetrically around zero, but it is much wider than the one for the Π copula. Moreover, the test statistic distributions are very different, as is seen from Figure 3.3(b). Although, during the simulations we observe test values coming from the flat dashed density (mixture of Frank I), we compare them with the density for the Π copula (the dotted short-dashed line). When the true copula is a Frank(0.5), the density of the test statistic even gets more concentrated around zero, and takes on values closer to zero for parameter values larger than 0.5. The situation for the mixture of Frank II is comparable to that for the mixture of Frank I. Here the density of the parameter estimates lies even more to the right than for Frank(0.5), but the test statistic density stays flat in comparison to the one coming from Frank(0.5). This explains the high power results of the parametric models in the examples of the mixture of Frank copulas.

The other observation is that for the asymmetric copula model the density of the parameter estimate is very similar to the one for the Π copula, but shifted slightly to the negative side. There is still a high percentage of positive parameter estimates as in the example of mixture of Frank I, yet the test statistic density is again very different. It is narrower and concentrated closer around zero. This explains the discrepancy in power between the two examples.



(a)



(b)

Figure 3.3: *(a). Densities of the parameter estimates ($\mathcal{C}$ = Frank); (b). Densities of AD test statistics (with indicated* 0.95 *quantile under* Π *copula).*

Size study

Keeping the size of the test occurs to be impossible for some PQD copulas in the considered PQD testing setting. Indeed, any of the described test statistics converges to zero (degenerately) when the true copula lies entirely above the Π copula. One could consider PQD copulas which have a common (measurably non-zero) part with the Π copula or a subset of PQD copulas "asymptotically close" to the Π copula, e.g., $C - \Pi \geq o\big(n^{-\alpha}\big)$. Both of these considerations are beyond the scope of this chapter.

Nevertheless, we present finite-sample performances for two examples of PQD copulas coming from the Frank and Clayton families with theoretical Kendall's tau equal to 0.05. Table 3.5 contains results for sample sizes 200, 1000 and 3000. The results for the non-parametric and Π-reference approaches for the larger sample sizes were all equal to zero, and hence are not included in the table. Moreover, as the Clayton copula has heavier tails than the Frank one, we see that the results when using the parametric reference model $\mathcal{C}$ = Clayton are smaller than the results for the model $\mathcal{C}$ = Frank.

Lastly, Table 3.4 shows that all of the testing approaches approximately hold the level for finite-sample sizes when the true copula is the Π copula.

distance	non-parametric	Π-reference	parametric $\mathcal{C}=$ Frank	parametric $\mathcal{C}=$ Clayton
		$n = 200$		
KS	0.047	0.042	0.045	0.044
CvM	0.046	0.050	0.050	0.050
AD	0.050	0.050	0.050	0.050
		$n = 1000$		
KS	0.057	0.045	0.048	0.048
CvM	0.053	0.056	0.057	0.056
AD	0.066	0.056	0.057	0.057

Table 3.4: *Size study results for the* Π *copula.*

distance	$n = 200$ non-parametric	$n = 200$ Π-reference	$n = 200$ parametric $\mathcal{C} =$ Frank	$n = 200$ parametric $\mathcal{C} =$ Clayton	$n = 1000$ parametric $\mathcal{C} =$ Frank	$n = 1000$ parametric $\mathcal{C} =$ Clayton	$n = 3000$ parametric $\mathcal{C} =$ Frank	$n = 3000$ parametric $\mathcal{C} =$ Clayton
			$C =$ Frank(0.5)					
KS	0.011	0.008	0.014	0.013	0.020	0.004	0.033	0.007
CvM	0.010	0.010	0.011	0.011	0.005	0	0.024	0.004
AD	0.009	0.007	0.012	0.012	0.014	0.011	0.030	0.019
			$C =$ Clayton(0.11)					
KS	0.006	0.003	0.011	0.007	0.024	0.006	0.091	0.019
CvM	0.004	0.003	0.003	0.003	0.009	0.002	0.057	0.013
AD	0.004	0.004	0.004	0.004	0.011	0.008	0.041	0.021

Table 3.5: *Size study results for a selection of copula families.*

3.5 Danish fire insurance data

The Danish fire insurance data set consists of 2167 claims over 1 million Danish Krone (DKK) on fire insurance in Denmark in the years 1980 to 1990. There are three types of claims referring to losses in buildings (B), their content (C) and profit they generated (P). The multivariate Danish fire insurance data are publicly available at `http://www.ma.hw.ac.uk/~mcneil/data.html`. In a univariate setting the sum of three variables has been studied in Embrechts et al. (1997), McNeil (1997) and Scaillet (2004).

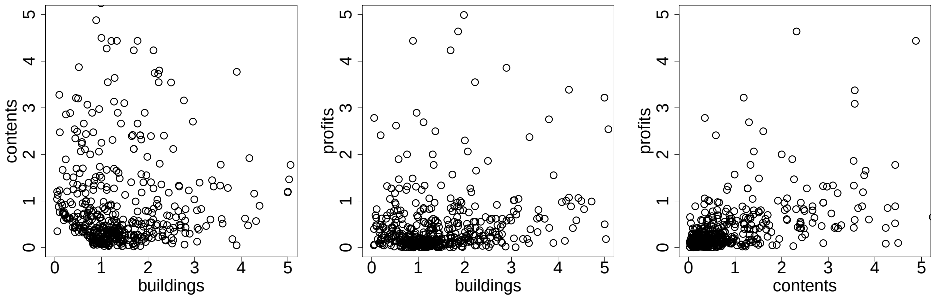
We consider only positive claims in all three variables, which reduces the sample size to 517. Figure 3.4 provides pairwise scatter plots of the observations (a) (restricted to the $[0, 5]^2$ region, for a good visual impression of the relations in the observed data) and all pseudo-observations (b). Figure 3.4 (c) presents the level plots of the estimated copula function minus the independence copula function. The violation regions, colored in dark green, are clearly visible in the left plot. They are rather small in the middle plot and are not present in the right plot at all. Moreover, Table 3.6 contains summary statistics of this reduced data set and we can see that the data are highly positively correlated, but to answer the question if a pair is PQD we employ the proposed testing procedures.

	building	content	profit
Min.	0.0482	0.0254	0.0041
1st Qu.	0.8377	0.2903	0.0964
Median	1.2816	0.6030	0.2451
Mean	2.1592	2.3712	0.8572
3rd Qu.	2.1104	1.5241	0.6401
Max.	95.1684	132.0132	61.9327

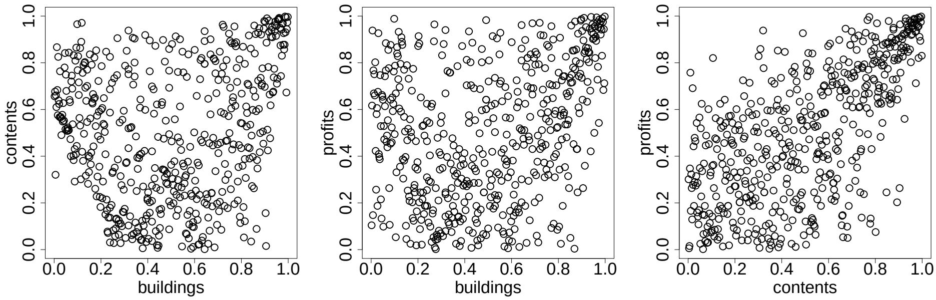
Table 3.6: *Summary statistics for the Danish fire insurance data.*

Table 3.7 contains estimates of Pearson's correlation coefficient, Kendall's tau and the univariate parameter of the Frank copula family, revealing that the dependence is strongest between the variables content and profit.

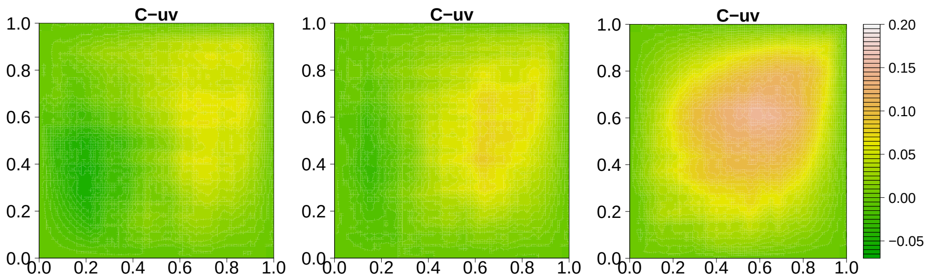
For illustration purposes and according to our recommendations from the simulation section, we present results only for the AD-based tests. The p-values in Table 3.8 suggest that we strongly reject the null hypothesis of PQD for the (B, C) pair for all the tests. It is most likely the effect of the gap in (pseudo-)data in the lower-left range region. The estimate of the copula in this region is zero, thus provides the most substantial violations of the PQD



(a)



(b)



(c)

Figure 3.4: *Danish fire insurance data. (a): observations (restricted to* $[0,5]^2$ *region); (b): pseudo-observations; (c) level plots of estimated* $C - \Pi$*.*

condition. Moreover, next to the concentration of the pseudo-observations towards the upper-right corner of the unit square, there is a visible increased denseness towards the upper-left direction, which also proves against PQD.

	(B,C)	(B,P)	(C,P)
Pearson's cor.coef.	0.6269	0.7910	0.6174
Kendall's tau	0.1172	0.2009	0.4620
$\hat{\theta}_{\tau_n}$	1.0669	1.8699	5.0853

Table 3.7: *Estimates of selected parameters.*

	(B,C)	(B,P)	(B,C) $u,v > 0.2$
non-parametric	0	0.0217	0.7158
Π-reference	0.0004	0.2706	0.9998
parametric ($\mathcal{C}$=Frank)	0	0	0.9997

Table 3.8: *Estimated p-values for the proposed tests based on the AD-distance.*

Concerning the second pair of variables (B,P) there is not as much agreement between the results of the three tests. The Π-reference test gives no evidence to reject the null hypothesis and the other two tests reject it. In the scatter plot of the pseudo-data for this pair of variables we also see a concentration of points towards the upper-left corner as in the first pair. This is probably enough for the non-parametric method to catch it. For the parametric method this effect is even amplified by a high value of the empirical Kendall's tau.

For the third pair (C,P) the test statistic occurs to be equal to zero, thus any distribution of the test statistic is uninformative and thus the resampling process is unnecessary. We naturally do not reject the null hypothesis.

It occurs that the (B,C) and (C,P) pairs are presumably easily decided upon. The (B,P) pair, although much more PQD-looking, still does not follow a PQD structure in our view.

The PQD structure is a global feature of a copula. Specifically in the considered example, practitioners might be interested in testing for a local shape of the structure, e.g., for the (B,C) pair the underlying copula might lie significantly above the Π copula in the region where (for example) $u,v > 0.2$. Our tests can be easily modified to check that, simply by restricting the integration in the test statistic to the particular local region, for example $(u,v) \in (0.2,1]^2$. We carried out this local tests for (B,C) pair and report on the resulting p-values in the last column of Table 3.8. Note, that restricting the region results in not rejecting the null hypothesis by any of the methods. This result nicely complement our former analysis of the (B,C) pair on whole the region.

Finally, we mention that the PQD condition extends to higher dimensions $(X_1, \dots, X_d)$ in a form of

- Positive Lower Orthant Dependence (PLOD)

$$\forall x_1, \dots, x_d \qquad \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d) \geq \prod_{i=1}^{d} \mathbb{P}(X_i \leq x_i),$$

- Positive Upper Orthant Dependence (PUOD)

$$\forall x_1, \dots, x_d \qquad \mathbb{P}(X_1 > x_1, \dots, X_d > x_d) \geq \prod_{i=1}^{d} \mathbb{P}(X_i > x_i),$$

- Positive Orthant Dependence (POD): when both PLOD and PUOD hold.

In higher dimensions one is, of course, facing the curse of dimensionality: there is a need for larger data sets to fill the gaps in the space. Apart from this, it is straightforward to extend the Π-reference and the parametric reference model approach to higher dimensions. The extension of the proposed non-parametric approach to higher dimensions is more tricky, since the method requires computing higher order partial derivatives, which are necessary to estimate the conditional distribution functions, i.e., similarly as in Theorem 1.2,

$$U_2 = C_{U_1}^{-1}(T_1) \quad \& \quad U_3 = \left(\frac{C_{U_1,U_2}(\cdot)}{C_{U_1,U_2}(1)}\right)^{-1}(T_2) \quad \Longrightarrow \quad (U_1, U_2, U_3) \sim C,$$

where $(U_1, T_1, T_2) \sim_{iid} U[0,1]$ and $c_{u_1,u_2}(u_3)$ denotes the second order partial derivative of $C(u_1, u_2, u_3)$ with respect to u_1 and u_2. This demands extensive computational resources. Table 3.9 contains p-values of the tests applied to all the 3 variables. We can see that the Π-reference and the parametric approaches lead to different conclusions. The non-parametric approach rejects both PLOD and PUOD, whereas the Π-reference approach rejects none of the two. Given the revealed violations of PQD-ness in for example the pair (B, C) one would expect at least one of the two conditions (PLOD or PUOD) to be rejected. The trivariate Frank copula only depends on one parameter and its appropriateness here is even more of an issue than in the previous bivariate case.

	PLOD	PUOD
non-parametric	0.0274	0.0011
Π-reference	0.7318	0.2143
parametric ($\mathcal{C}$=Frank)	0	0.0014

Table 3.9: *Estimated p-values based on AD-distance for the 3 variables.*

3.6 Conclusions and further discussion

In this chapter we proposed a constrained copula estimation approach to the PQD testing problem. Via constrained copula estimation and resampling from this estimated distribution, we assess the distribution of the test statistic under the null hypothesis. In the case of non-parametric PQD-constrained copula estimation, a monotonic rearrangement technique was used to be able to carry out the resampling. Similar techniques of monotonic rearrangements have, in the statistics literature, been used in estimation of unimodal densities by Fougères (1997), or monotone regression functions in Dette et al. (2006) and, in estimation of regression quantiles in Dette and Volgushev (2008) and Chernozhukov et al. (2009). Note that in our copula context the monotonic rearrangement technique is used on a partial derivative of a bivariate function. Theoretical studies of the final estimates are quite involved here, and are part of future research.

The finite-sample performances of the constrained copula estimation built testing procedures were compared to the finite-sample performance of the Π-reference approach, available in the literature. The proposed testing procedures always outperform the Π-reference approach.

The main idea employed was to explore the local structure of a copula estimator, as the PQD condition can be violated on a subset of the unit square. Specifically, if $C > \Pi$ on a non-zero measure subset A of the unit interval, then $\mathbb{P}\big(\Pi(u,v) - C_n(u,v) < 0 : \ (u,v) \in A\big) \to 1$, hence the test statistic under C should be smaller than the one under Π, leading to smaller critical values. Smaller critical values under the alternative lead to increase in power. At the same time, the new procedures manage to hold the level of the test for the Π copula.

Eventually, we tested a real data example of Danish fire insurance data for PQD. Within the context of this example we also discussed the extendibility of the methods to higher dimensions and mentioned the challenges related to this.

In the next chapter we investigate tail monotonic properties in testing for another dependence structures.

Chapter 4

Testing tail monotonicity by constrained copula estimation

4.1 Introduction

This chapter is based on Gijbels and Sznajder (2011c) and develops tests for tail monotonicity.

Tail monotonicity was among the dependence concepts discussed in Lehmann (1966) and was further investigated in Esary and Proschan (1972). For two random variables X and Y, Y is said to be left tail decreasing (LTD) in X if and only if

$$\mathbb{P}(Y \leq y | X \leq x) \quad \text{is non-increasing in } x \quad \forall y \; .$$

This type of dependency is on the one hand stronger than X and Y being positive quadrant dependent (PQD) which holds if and only if

$$\mathbb{P}(Y \leq y, X \leq x) \geq \mathbb{P}(Y \leq y)\mathbb{P}(X \leq x) \quad \forall x, y \; , \tag{4.1}$$

and is on the other hand weaker than Y being positively regression dependent on X which holds if and only if

$$\mathbb{P}(Y \leq y | X = x) \quad \text{is non-increasing in } x \quad \forall y \; .$$

The notion of positive regression dependency goes back to Tukey (1958) and Lehmann (1959). Esary and Proschan (1972) showed that Y being left tail decreasing in X implies that $\mathrm{Cov}(f(X,Y), g(X,Y)) \geq 0$ for all non-decreasing functions f and g. The above relationships between these positive dependence structures were pointed out in Lehmann (1966). Parallel definitions of similar notions of negative dependence structures can be given. The bivariate positive dependence concepts of left tail decreasingness and right tail increasingness have been generalized to positive dependence orderings in Colangelo (2008). See also Colangelo et al. (2005) and Colangelo et al. (2006) for studies on multivariate positive dependencies.

Investigating the specific dependence structure between random variables is of great importance in many area's. The study of tail dependencies, and in particular testing whether a specific tail monotonicity dependency structure holds or not, is of interest for applications in insurance and finance, among others. Finding out relations between tail distributions corresponding to, e.g., large losses and large claims is a main concern for insurance companies.

Testing for positive quadrant dependence has received considerable attention in the recent literature including the papers of Denuit and Scaillet (2004), Scaillet (2005), and Gijbels et al. (2010) and Gijbels and Sznajder (2011a), among others. Up to the authors' knowledge, tail monotonicity has not yet been a subject for testing in statistics. It might however be of interest to practitioners as it gives more insight in the overall conditional distribution or survival functions, and it allows for more flexibility than regression monotonic structures. In the application section we apply tail monotonicity tests to data examples from insurance, finance and ecological studies.

The contribution of this chapter consists of developing testing procedures for the following testing problem

$$\begin{aligned} &H_0: \ Y \text{ is left tail decreasing in } X \\ &\text{versus} \\ &H_1: \ Y \text{ is not left tail decreasing in } X\ . \end{aligned} \tag{4.2}$$

When as a result of a test a positive quadrant dependence structure has not been rejected, it is worthwhile to explore further and test whether the more stringent positive dependence structure "Y is left tail decreasing in X" holds or not. In the first step we reformulate testing problem (4.2) in terms of copulas. The basic idea is then to look at test statistics that describe the discrepancy between a non-constrained copula estimator and an LTD-constrained copula estimator. Major issues when constructing a constrained copula estimator are: (i) this estimator should have the properties of a bivariate distribution function; (ii) a method of how to resample from it needs to be worked out. Both issues require the use of innovative techniques. Similar techniques were applied in the previous chapter in the simpler context of testing for positive quadrant dependence. Although the focus is on deriving a testing procedure for the testing problem (4.2), the developed methodology easily applies to testing for other type of tail monotonicity structures.

The chapter is organized as follows. In Section 4.2 we briefly recall the definitions of tail monotonic structures, their properties and connections with quadrant dependency. In Section 4.3 we develop the testing procedure and define the constrained copula estimator. Section 4.4 then discusses approaches to assess the distribution of the test statistic under the null hypothesis. Monte Carlo power simulation results are gathered in Section 4.5 and in Section 4.6 we apply the testing procedures to real data examples. Some conclusion of the contributions in this chapter are provided in Section 4.7.

4.2 Tail monotonicity

Tail monotonicity as defined in Esary and Proschan (1972) describes the monotonic behaviour of a tail (left or right) of the "conditional" distribution of Y given X. Similar to the definitions of LTD and RTI in Definition 1.13 one can define left tail increasing (LTI) and right tail decreasing (RTD) dependence structures. Note from Definition 1.13 that tail monotonicity is not a symmetric concept, and thus the conditional distributions of $Y|X$ and $X|Y$ have to be treated separately. This is in contrast to positive quadrant dependence, where the role of X and Y can be interchanged and hence is a symmetric concept.

The reformulation of Definition 1.13 in terms of the copula function in Proposition 1.12 will be exploited in the sequel to derive testing procedures for tail monotonicity. We develop a test for testing the null hypothesis that Y is left tail decreasing in X. However, exploiting the relations among the different tail dependence concepts, as stated in Proposition 4.1, easily extends to testing procedures for any other specific monotonic tail dependence structure. Proposition 4.1 shows how the various dependence concepts are related among themselves under monotonic transformations of the marginal distributions.

Proposition 4.1. *Let α be a strictly increasing real function, and β_1 and β_2 be strictly decreasing real functions. Then*

(a) $\alpha(Y)|\beta_1(X)$ is left tail decreasing if and only if $Y|X$ is right tail decreasing, i.e.,

$$LTD\,(\alpha(Y)|\beta_1(X)) \Longleftrightarrow RTD(Y|X)$$

(b) $\beta_2(Y)|\alpha(X)$ *is left tail decreasing if and only if* $Y|X$ *is left tail increasing, i.e.,*

$$LTD\left(\beta_2(Y)|\alpha(X)\right) \Longleftrightarrow LTI(Y|X)$$

(c) $\beta_2(Y)|\beta_1(X)$ *is left tail decreasing if and only if* $Y|X$ *is right tail increasing, i.e.,*

$$LTD\left(\beta_2(Y)|\beta_1(X)\right) \Longleftrightarrow RTI(Y|X).$$

So when the interest is in testing for "Y is right tail increasing (RTI) in X" it thus simply suffices to consider, for example, $\beta_1(x) = \beta_2(x) = -x$ and hence the couple $(-X, -Y)$, and apply the testing procedure for left tail decreasingness to this couple. By similar considerations we can test for any other specific tail monotonic dependency structure.

To prove Proposition 4.1 we first need to see how the copula of a transformed couple of random variables, say $(\alpha(X), \beta(Y))$, with $\alpha(\cdot)$ and $\beta(\cdot)$ being monotonic transformations, relates to the copula of the original couple (X, Y). This information is provided in the following lemma (see Theorem 2.4.4. in Nelsen (2006)).

Lemma 4.1. *Let* α *be a strictly increasing real function, and* β_1 *and* β_2 *be strictly decreasing real functions. Then*

(a) $C_{\alpha(X),\beta_2(Y)}(u, v) = u - C_{X,Y}(u, 1 - v)$

(b) $C_{\beta_1(X),\alpha(Y)}(u, v) = v - C_{X,Y}(1 - u, v)$

(c) $C_{\beta_1(X),\beta_2(Y)}(u, v) = u + v - 1 + C_{X,Y}(1 - u, 1 - v).$

We now provide the proof of Proposition 4.1.

Proof of Proposition 4.1.

We only explicitate the proof of item (c) as the methodology of proof is the same for the other items.

The random variable $\beta_2(Y)$ is left tail decreasing in $\beta_1(X)$ if and only if $C_{\beta_1(X),\beta_2(Y)}(u, v)/u$ is non-increasing in u for any fixed v. According to Lemma 4.1 (c) we have that

$$\frac{C_{\beta_1(X),\beta_2(Y)}(u, v)}{u} = \frac{u + v - 1 + C_{X,Y}(1 - u, 1 - v)}{u}\,.$$

The right-hand side of the last equality being non-increasing in u is equivalent to

$$\frac{1-u+1-v-1+C_{X,Y}(u,v)}{1-u}$$

being non-decreasing in u for any fixed v, which is exactly as the copula definition of RTI in Proposition 1.12. $\square$.

In the dependence structure world LTD is a form of a positive dependence, that is a stronger and more constrained relation than positive quadrant dependence. In particular according to Nelsen (2006)

Proposition 4.2. *If $X|Y$ or $Y|X$ are left tail decreasing or right tail increasing, then (X,Y) are positive quadrant dependent, i.e.,*

$$[LTD(Y|X) \text{ or } LTD(X|Y) \text{ or } RTI(Y|X) \text{ or } RTI(X|Y)] \Longrightarrow PQD(X,Y).$$

We shall use C to denote $C_{X,Y}$ or $C_{U,V}$ (where $U = F(X)$ and $V = G(Y)$) and LTD for $LTD(Y|X)$ or $LTD(V|U)$, unless the more detailed notations are necessary.

4.3 LTD adjustment and test statistic

In this section we discuss how to develop a testing procedure for the testing problem (4.2). Exploiting the equivalent copula function condition for LTD in (1.11) and denoting

$$C_v(u) = \frac{C(u,v)}{u} \tag{4.3}$$

we can reformulate testing problem (4.2) as follows:

$H_0:\ C_v(u)$ is non-increasing in u for every v

versus

$H_1:\ C_v(u)$ is not non-increasing in u for some v .

4.3.1 LTD adjustment

Suppose for a moment that the copula function C would be known, and we would like to measure how far the copula is away from a "closest" LTD-copula. Of course we then need to define a measure of closeness, as well as an LTD-constrained copula constructed from the given copula C. In other words, for a given copula C (not necessarily LTD) we are searching for an LTD-constrained copula $\widetilde{\widetilde{C}}$ which is closest in some sense to C.

We first describe how to deal with the second issue. First of all we need to impose condition (1.11) on that given copula C. Therefore, we construct its "non-increasing hull" as a bivariate function

$$\overline{C}(u,v) = u \max_{u \leq t \leq 1} C_v(t) \; . \tag{4.4}$$

Obviously the function

$$\overline{C}_v(u) = \frac{\overline{C}(u,v)}{u} . \tag{4.5}$$

is a non-increasing function in u, for all fixed v. As an example we depict in Figure 4.1 the copula function A, defined in (4.18) in Section 4.5, together with its corresponding "non-increasing hull" function $\overline{C}(u,v)$. It can be clearly seen from Figures 4.1 (c) and(d) that the function $C_v(u)$ violates the non-increasingness property for certain values of v, whereas the function $\overline{C}_v(u)$ is non-increasing as a function of u for all values of v. This is in particular visible from the lower left corner in the graphs.

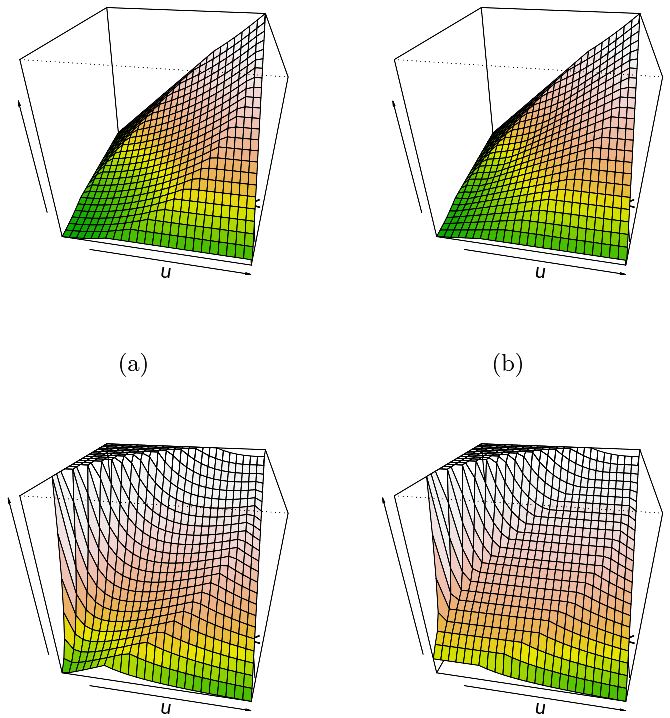
Note from the definitions that if $C_v(u)$ is non-increasing for a fixed v, then $\overline{C}(u,v) \equiv C(u,v)$ as a function of u. Further note that $\overline{C}_v(u) \geq C_v(u)$, for all u and v.

For a given copula C, its "closeness" to the LTD-constrained bivariate function $\overline{C}$ is then measured via

$$\int \left(\overline{C}_v(u) - C_v(u)\right) dudv \; . \tag{4.6}$$

In case of an unknown copula, the functions $\overline{C}_v(u)$ and $C_v(u)$ need to be estimated from the available data as we describe later on in this section.

For assessing the distribution of a test statistic based on an empirical version of (4.6) we need to be able to resample under H_0, and hence need to be able to resample from an LTD-constrained copula estimator. This still requires more work. Indeed, $\overline{C}$ in (4.4) is a bivariate function on the unit square, but in general it is not a distribution function. However, we can create one from $\overline{C}$ by applying a monotone rearrangement technique to its partial derivative. Definition 1.7 and Proposition 1.7 provide us with the tools to continue. Indeed, from items (b) and (d) in Proposition 1.7 it is clear that if we have a function f that is possibly not non-decreasing we can force it to be non-decreasing, such that this forced non-decreasing version is



(a) (b)

(c) (d)

Figure 4.1: *Illustration for the copula A: (a) the copula A; (b) its corresponding 'non-increasing hull' function $\overline{C}(u,v)$; (c) the function $C_v(u)$ and (d) the function $\overline{C}_v(u)$.*

closest in L_p-sense to the original not non-decreasing version. Furthermore, item (c) tells us that if the original function is already non-decreasing, then this operation alters nothing to it. In addition, item (a) tells us that the operation is unique.

We apply this monotonic rearrangement technique to obtain a proper bivariate distribution function that satisfies the LTD-constraint. Therefore we apply the technique to the function

$$\bar{c}_u(v) = \frac{\partial \overline{C}(u,v)}{\partial u} \; , \tag{4.7}$$

and define a new (proper) bivariate distribution function as follows

$$\widetilde{C}(u,v) = \int_0^u \widetilde{c}_t(v)dt \ , \tag{4.8}$$

where

$$\widetilde{c}_u(v) = (\xi \circ \xi)\left(\overline{c}_u(\cdot)\right)(v) \ . \tag{4.9}$$

In Figure 4.2 we depict for the copula A the function $\overline{c}_u(v)$ which, for this example, is a non-decreasing function in v for all fixed u, and hence is not altered by the monotonic rearrangement operation.

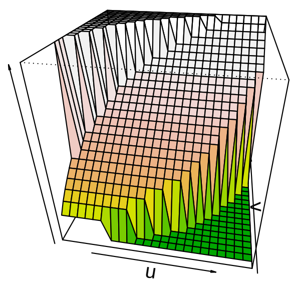


Figure 4.2: *Illustration for the copula A: the function $\overline{c}_u(v)$.*

As $\widetilde{C}$ is a cumulative distribution function on the unit square with continuous margins, there exists its unique copula function (denoted by $\widetilde{\widetilde{C}}$), i.e., if $(U,\widetilde{V}) \sim \widetilde{C}$, then

$$\widetilde{\widetilde{C}}(u,v) = \widetilde{C}\left(u, G_{\widetilde{V}}^{-1}(v)\right) \ ,$$

where $G_{\widetilde{V}}^{-1}$ is a pseudo-inverse of the cumulative distribution function of $\widetilde{V}$, and $G_{\widetilde{V}}(v) = \widetilde{\widetilde{C}}(1,v)$.

Finally, note that for $(U,\widetilde{V}) \sim \widetilde{C}$, with $U \sim \mathcal{U}[0,1]$ distributed random variable, we have

$$\mathbb{P}(\widetilde{V} \leq v | U = u) = \frac{\partial \widetilde{C}(u,v)}{\partial u} = \widetilde{c}_u(v) \ . \tag{4.10}$$

This will be the key for the random number generation process in Section 4.4.

4.3.2 Test statistic

Our test statistic will be based on an empirical version of (4.6). Given an i.i.d. sample $\{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ from (X, Y) we estimate the unknown copula function $C(\cdot, \cdot)$ by the empirical copula estimator of Deheuvels (1979). The corresponding estimates of the functions $C_u(v)$ and $\overline{C}_u(v)$, defined in respectively (4.3) and (4.5), are then given by

$$C_{v,n}(u) = \frac{C_n(u,v)}{u}$$

and

$$\overline{C}_n(u,v) = u \max_{u \le t \le 1} C_{v,n}(t) \qquad \text{and} \qquad \overline{C}_{v,n}(u) = \max_{u \le t \le 1} C_{v,n}(t).$$

Inspired by (4.6), we then consider the test statistic

$$S_n = \sqrt{n} \int \left(\overline{C}_{v,n}(u) - C_{v,n}(u)\right) dudv = \sqrt{n} \int \frac{\overline{C}_n(u,v) - C_n(u,v)}{u} dudv \, . \tag{4.11}$$

The asymptotic variance of the empirical copula $C_{v,n}(u)$ is given by

$$\frac{C\big(1-C\big) - 2c_u(v)C(1-u) - 2c_v(u)C(1-v)}{nu^2} + \frac{c_u(v)c_v(u)\big(C-uv\big) + c_u^2(v)u(1-u) + c_v^2(u)v(1-v)}{nu^2} \, ,$$

which can be unbounded for fixed n and u close to zero. Therefore, we propose to limit the integration region in the test statistic (4.11) to $(n^{-1/2}, 1] \times [0, 1]$ for $u \times v$.

4.4 Assessing the distribution of the test statistic under the null hypothesis

To obtain critical values for the test we propose a finite sample approach by means of resampling under the null hypothesis. We compare approaches based on: (i) resampling from an estimated non-parametric distribution and (ii) resampling from a reference type of a parametric copula family. Approaches similar to these have been used in the previous chapters in the simpler context of testing for positive quadrant dependence.

Nonparametric resampling

The non-parametric approach is based on resampling from a smooth estimate of (4.8). This is obtained by looking at the smoothing problem in a bivariate regression context, with $\overline{C}_n$ the response variable, using the bivariate local linear regression smoothing technique, as described in Section 1.2.3. Obviously any other bivariate regression smoothing technique might be used.

We consider the $c_{i,j}$'s in the regression setting (1.4) to be equal to $\overline{C}_n$ values on the fixed grid of points $\{(u_i, v_j)\}_{i,j=0}^{m+1}$, i.e.,

$$c_{i,j} = \overline{C}_n(u_i, v_j) = u_i \max_{i \le \ell \le m+1} \frac{C_n(u_\ell, v_j)}{u_\ell} \, . \tag{4.12}$$

The local linear estimate of (4.7) is then given by $c_{u,n}(v)$ from (1.4). Furthermore, let

$$\widetilde{c}_{u,n}(v) = (\xi \circ \xi)\left(\bar{c}_{u,n}(\cdot)\right)(v) \tag{4.13}$$

be an estimator of (4.9). Note, that ξ is an operator on a space of functions, yet we can approximate it with prescribed precision on a finite (dense enough) sequence of points. Also note, that because of the finite sample error, (4.13) is likely not to be a valid cumulative distribution function (contrary to (4.9)), so we transform it linearly into one

$$\widehat{\widetilde{c}}_{u,n}(v) = \frac{\widetilde{c}_{u,n}(v) - \widetilde{c}_{u,n}(0)}{\widetilde{c}_{u,n}(1) - \widetilde{c}_{u,n}(0)}. \tag{4.14}$$

Eventually, (4.14) is a valid (random) univariate cumulative distribution function and an estimator of (4.9), which is also a partial derivative of (4.8), and hence an estimator of the conditional distribution function of $\widetilde{V}|U$ (see (4.10)), which can be used to resample from the estimated (4.8).

Indeed, for a fixed sample (corresponding to a random event ω) let us sample two sets of observations from the uniform $[0, 1]$ distribution $\{u_i^*\}_{i=1}^n$ and $\{t_i^*\}_{i=1}^n$, and compute

$$\widetilde{v}_i^* = \left(\widehat{\widetilde{c}}_{u_i^*,n}^{\,-1}(\omega)\right)(t_i^*)$$

to obtain a sample $\{u_i^*, \widetilde{v}_i^*\}_{i=1}^n$ coming from the estimated (4.8). If we repeat the resampling N times and at each step compute the realized test statistic (4.11) $S_{n,j}^*(\omega)$, we obtain an approximated distribution of the test statistic under (4.8), which can be used to obtain an approximate p-value of the test for a given sample

$$p^*(\omega) = \frac{1}{N}\sum_{j=1}^{N} \mathbb{I}\left(S^*_{n,j}(\omega) > S_n(\omega)\right). \tag{4.15}$$

As will be seen from the simulation study in Section 4.5, this non-parametric approach tends to exceed the prescribed nominal significance level in the independence copula example. This tendency to exceed the nominal level is caused by the fact that we resample each time from a copula which is a bit “too LTD” as we use the empirical non-increasing envelope (4.12) as a starting point for obtaining a smooth derivative estimate. This can be improved upon by using a less crude way to force the empirical copula C_n to be such that $C_n(u,v)/u$ is a non-increasing function for each fixed v. This can for example be achieved by applying a constrained least squares fit instead. Specifically, we apply the cone projection algorithm described in Meyer (2008) to the set of values $\left\{\frac{C_n(u_i,v_j)}{u_i} - v_j\right\}_{i=0}^{m+1}$, and this for each fixed v_j, and we denote the obtained copula values as $\{C_n^{\mathrm{LSC}}(u_i,v_j)\}_{i=0}^{m+1}$. The iterative algorithm of Meyer (2008) allows for solving a general problem of the following form:

$$\hat{\zeta} = \mathrm{argmin}||\mathbf{x} - \zeta||^2 \quad \text{such that} \quad A\zeta \geq 0, \tag{4.16}$$

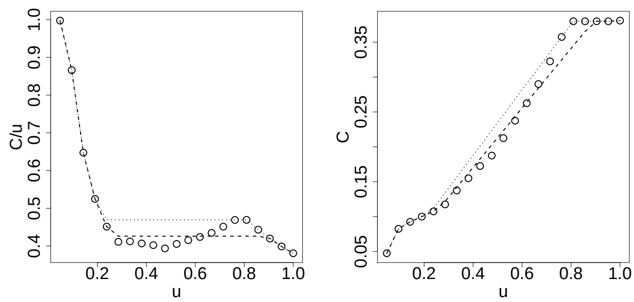
where $\mathbf{x}$ is a vector of observations and A is an irreducible matrix of linear conditions. In our case

$$A = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 & -1 \\ 1 & -1 & 0 & \dots & 0 & 0 & 0 \\ & \ddots & & \ddots & & \ddots & \\ 0 & 0 & 0 & \dots & 0 & 1 & -1 \end{bmatrix},$$

where the first two rows correspond to the boundary condition $C(1,v) = v$ and the next rows enforce non-increasingness. The implemented algorithm is available at `http://www.stat.colostate.edu/~meyer`.

Now, instead of using the matrix Y for smoothing in (1.4), we use

$$Y^{\mathrm{LSC}} = \begin{bmatrix} C_n^{\mathrm{LSC}}(u_0,v_0) \\ \vdots \\ C_n^{\mathrm{LSC}}(u_{m+1},v_{m+1}) \end{bmatrix}. \tag{4.17}$$



(a) (b)

Figure 4.3: *Illustration of copula adjustment for the Copula A in (4.18): (a) empirical copula C_n section for fixed v (points), corresponding C_n^{LSC} (dashed line) and $\overline{C}_n$ (dotted line) values; (b) the same plot on the copula level.*

The difference between the two approaches is illustrated in Figure 4.3 for a simulated sample from copula A in (4.18).

In Section 4.5 we investigate the finite-sample performance of the resulting two non-parametric methods:

Nonparametric method I: using $\overline{C}_n$ in (4.12) to enforce the non-increasingness assumption;

Nonparametric method II: using the less crude C_n^{LSC} to enforce this assumption.

Resampling by making reference to a parametric copula family

We need to resample under the null hypothesis. Instead of relying on the non-parametric LTD-constrained estimator one can also resample from a parametrically estimated LTD-copula. Essential is then to consider a parametric family of copulas that results into an LTD-copula when restricting the parameter space.

An extreme case within this approach is to simply resample from the independence copula Π, since this copula obviously is LTD. We refer to this approach as the Π-reference approach.

Note that the independence copula is often a boundary point between positive and negative dependencies in parametric families of copulas. This is also the case in tail monotonicity considerations as the independent distribution of (X, Y) is at the same time left and right tail increasing and decreasing. Therefore, it is of interest to consider resampling (N times) from the independent distribution, and obtain an approximate p-value in the same way as in (4.15).

Naturally, one can focus on any another fixed distribution if it is a priori suitable in a particular study. However, the resampling procedure entirely depends on that fixed parametric choice. A generalization of such an approach is to consider a parametric copula family as a reference for the resampling purpose. Such a parametric family of copulas should at least include a subset of copulas that are LTD. For brevity of presentation we restrict the simulation study in Section 4.5 to the same copula families considered in previous chapters, namely Frank and Clayton one parametric Archimedean copula families.

In this family a positive parameter θ results in an LTD-copula whereas a negative value of θ gives a copula that is not LTD. Moreover, there exists a bijection between the parameter θ and Kendall's tau, and this relation is used in the estimation procedure (see Genest (1987)). The constrained parameter estimation in this case relies on projecting an estimator (see Fils-Villetard et al. (2008)) on the positive half-line.

4.5 Simulation study

A Monte Carlo simulation study was conducted to investigate the finite-sample power performance of the testing procedures. According to Proposition 4.1 we can perform tests for all of the tail monotonic relations by applying the proposed testing procedure to appropriately transformed marginal observations, e.g., $\alpha(x) = x$ and $\beta_1(x) = \beta_2(x) = -x$.

In the simulation study we consider four true copulas: a Frank copula with parameter -1, the independence copula ($C = \Pi$), and two other copulas referred to as Copulas A and B respectively. The latter two copulas are elements of two broad collections of copulas described in Nelsen (2006). Copula A is a copula with mass distributed equally among two curves in the unit square, namely $u^2 + v^2 = 2u$ and $u^2 + v^2 = 2v$, and is defined by

$$C_A(u, v) = \min\left(u, v, \frac{u^2 + v^2}{2}\right) . \tag{4.18}$$

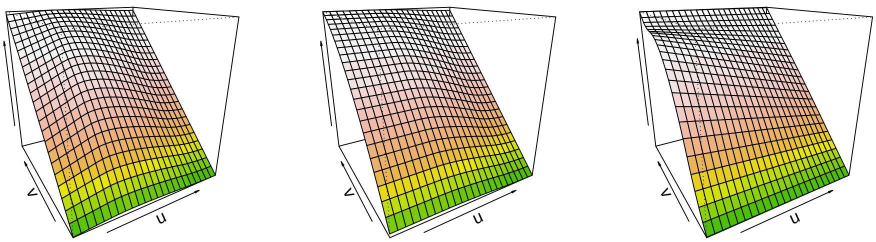
This copula possesses none of the tested eight dependence structures.

Copula B is an asymmetric absolutely continuous copula coming from a family of copulas with quadratic sections in v, as described in Section 1.2.2

in Proposition 1.6. In our example, ψ is a quadratic spline function

$$\psi(u) = 2u^2\mathbb{I}(0 \leq u \leq 0.25) + (-2u^2 + 2u - 0.25)\mathbb{I}(0.25 < u \leq 0.75) \\ + 2(1-u)^2\mathbb{I}(0.75 < u \leq 1).$$

This copula is neither symmetric in its tails nor in its variables, in particular it has no considered dependence properties for $Y|X$, but it is $LTD(X|Y)$ and $RTI(X|Y)$. As an illustration we plot in Figures 4.4 (a)—(b) the functions $C_u(v)$ and $\overline{C}_u(v)$, and in Figure 4.4 (c) the function $C_u(v)$ for the couple $(-V, -U)$. Recall that Copula B is not LTD for $Y|X$ and is RTI for $X|Y$.



(a) (b) (c)

Figure 4.4: *Illustration for the copula B in: (a) and (b) respectively the functions $C_v(u)$ and $\overline{C}_v(u)$ for the couple (U, V) and; (c) the function $C_v(u)$ for the couple $(-V, -U)$.*

These last simulation models are nontrivial cases of copulas which are PQD and not LTD. Furthermore, according to the results in Gijbels et al. (2010) and Gijbels and Sznajder (2011a) we can expect to wrongly reject approximately at most five percent of their samples when testing for PQD. Thus, the proposed LTD testing method allows to differentiate further between the two dependence structures.

Table 4.1 summarizes (indicated with an $\times$) the tail monotonicity dependencies that are present in each of the four copula models.

For each of the simulation models, we considered 1000 samples of sizes $n = 200$ and $n = 400$ to approximate the power. Every p-value was approximated by resampling 1000 times from the resampling distribution. The grid parameter was set to $m = \lfloor \sqrt{n} \rfloor$ and the bandwidth parameters h_i, $i = 1, 2$, were taken equal to $1.5/m$. We also used the same grid size for integrating the test statistic and applying the rearrangement operator to the estimated

	$Y\|X$				$X\|Y$			
copula	LTD	RTD	LTI	RTI	LTD	RTD	LTI	RTI
Frank(−1)		×	×			×	×	
Π-copula	×	×	×	×	×	×	×	×
Copula A								
Copula B					×			×

Table 4.1: *The copula simulation models and their dependence structures.*

partial derivatives. Finally, we rejected the null hypothesis for p-values lower than 0.05 (the considered significance level).

The simulation results are summarized in Tables 4.2 and 4.3, for respectively sample sizes $n = 200$ and $n = 400$. In the tables we present the proportion of times (over 1000 simulated samples) that the stated null hypothesis was rejected, and this for the testing procedures based on the three different approaches to the resampling as exposed in Section 4.4. We summarize results for the following five methods:

NON-PARAMETRIC I: based on $\overline{C}_n$ in (4.12);

NON-PARAMETRIC II: based on C_n^{LSC} and (4.17);

PARAMETRIC FRANK: using parametric resampling of Section 4.4 with as reference the Frank copula family;

PARAMETRIC CLAYTON: using parametric resampling of Section 4.4 with as reference the Clayton copula family;

Π-REFERENCE: using the independence copula for the resampling.

We first discuss the simulation results for sample size $n = 200$ reported in Table 4.2. For the Frank copula with parameter -1, all five methods work comparably well, with a slightly better performance for the non-parametric resampling methods. The results are similar for LTD and RTI in both tails as this copula is symmetric in its tails and in its variables. Moreover, the true negative dependence, in the form of LTI and RTD, is rarely rejected. Regarding this note however the different performance between the two parametric methods: when the parametric resampling is done from a Clayton reference family, the proportion of (wrong) rejections is around 14%. In other words, we get too many false rejections in these cases of negative dependencies.

		$Y\|X$				$X\|Y$			
copula	method	LTD	RTD	LTI	RTI	LTD	RTD	LTI	RTI
Frank(-1)	non-parametric I	0.617	0.012	0.014	0.635	0.613	0.021	0.012	0.616
	non-parametric II	0.590	0	0.002	0.616	0.591	0.001	0.001	0.593
	parametric Frank	0.586	0.025	0.022	0.610	0.581	0.036	0.035	0.580
	parametric Clayton	0.591	0.134	0.147	0.601	0.585	0.155	0.154	0.577
	Π-reference	0.590	0	0.002	0.612	0.592	0	0.001	0.585
Independence	non-parametric I	0.075	0.081	0.096	0.069	0.094	0.091	0.107	0.082
	non-parametric II	0.053	0.049	0.066	0.047	0.052	0.053	0.068	0.051
	parametric Frank	0.057	0.054	0.071	0.054	0.052	0.064	0.076	0.054
	parametric Clayton	0.064	0.073	0.082	0.070	0.068	0.077	0.092	0.074
	Π-reference	0.054	0.049	0.063	0.045	0.051	0.056	0.069	0.047
A	non-parametric I	0.973	1	1	0.955	0.946	1	1	0.961
	non-parametric II	0.883	1	1	0.858	0.865	1	1	0.886
	parametric Frank	0.997	1	1	0.995	0.995	1	1	0.996
	parametric Clayton	1	1	1	1	1	1	1	1
	Π-reference	0.022	1	1	0.010	0.010	1	1	0.024
B	non-parametric I	0.319	0.957	0.956	0.343	0.017	0.954	0.945	0.018
	non-parametric II	0.086	0.952	0.951	0.091	0	0.955	0.938	0.001
	parametric Frank	0.490	0.912	0.917	0.515	0.041	0.955	0.940	0.046
	parametric Clayton	0.859	0.907	0.921	0.880	0.323	0.959	0.938	0.315
	Π-reference	0.009	0.914	0.921	0.008	0	0.959	0.941	0

Table 4.2: *Simulation results for sample size $n = 200$. Proportions of rejections of the null hypothesis out of 1000 simulated samples.*

		$Y\|X$				$X\|Y$			
copula	method	LTD	RTD	LTI	RTI	LTD	RTD	LTI	RTI
Frank(-1)	non-parametric I	0.898	0.008	0.009	0.885	0.895	0.014	0.018	0.885
	non-parametric II	0.884	0.001	0.002	0.877	0.883	0	0	0.868
	parametric Frank	0.883	0.035	0.026	0.859	0.875	0.038	0.048	0.857
	parametric Clayton	0.882	0.246	0.252	0.863	0.882	0.288	0.306	0.852
	Π-reference	0.885	0.001	0.001	0.866	0.875	0	0	0.855
Independence	non-parametric I	0.083	0.086	0.090	0.093	0.081	0.083	0.104	0.083
	non-parametric II	0.057	0.050	0.049	0.054	0.047	0.048	0.060	0.048
	parametric Frank	0.059	0.054	0.056	0.062	0.049	0.058	0.064	0.053
	parametric Clayton	0.072	0.066	0.064	0.072	0.057	0.072	0.077	0.072
	Π-reference	0.053	0.047	0.049	0.049	0.041	0.047	0.058	0.047
A	non-parametric I	1	1	1	1	1	1	1	1
	non-parametric II	1	1	1	1	1	1	1	1
	parametric Frank	1	1	1	1	1	1	1	1
	parametric Clayton	1	1	1	1	1	1	1	1
	Π-reference	0.096	1	1	0.085	0.048	1	1	0.103
B	non-parametric I	0.559	1	1	0.585	0.017	0.999	1	0.010
	non-parametric II	0.232	1	1	0.240	0	0.998	0.999	0.002
	parametric Frank	0.819	0.996	0.997	0.836	0.058	0.999	0.999	0.059
	parametric Clayton	0.996	0.995	0.998	0.993	0.599	0.999	0.999	0.572
	Π-reference	0.031	0.996	0.997	0.029	0	0.999	1	0

Table 4.3: *Simulation results for sample size $n = 400$. Proportions of rejections of the null hypothesis out of 1000 simulated samples.*

If we look at the results for the independence copula the level is slightly higher than the prescribed five percent level for the non-parametric I method, and close to the five percent level for the non-parametric II method.

The situation for the singular copula A is correctly recognized in almost all of the cases except for the positive dependence in the Π-reference approach. This method seems to completely fail in this setting.

Concerning the results for the copula B, all of the methods correctly reject the majority of negative dependence structures. In the case of $Y|X$ and positive dependence the parametric Clayton method seems to work the best and the Π-reference has almost no power at all. In the reverse case $X|Y$ all but the parametric Clayton method correctly did not reject in most of the cases the positive dependence structures. In particular the results for this copula B simulation model show that the choice of the parametric copula family for resampling is an important issue. This choice is a clear drawback for such parametric resampling methods.

Recall that the non-parametric I method tends to reject more samples than the prescribed significance level of the test, and that the non-parametric II method is designed to correct for this. As a consequence, the non-parametric I method uniformly outperforms the non-parametric II method in terms of power in several copula simulation models.

The evolution of the results when the sample size increases from 200 to 400 can be seen from comparing Tables 4.2 and 4.3. For all four simulation examples we see that under the alternative the power increases for all of the methods. Under the independence model it fluctuates around the five percent level and in many cases comes closer to this nominal level. Moreover, the evolution of the results while going from sample size 200 to 400 for the copula B model, suggests that a 'bad' choice of the resampling parametric family may lead to inconsistency under the null hypothesis. Indeed, the rejection proportions for LTD in the $X|Y$ case increase from 0.323 to 0.599 for the parametric Clayton reference method, getting thus further away from the 0.05 level.

4.6 Real data examples

4.6.1 Danish fire insurance data

The Danish fire insurance data example introduced in Section 3.5 consists of observations on three random variables referring to insurance losses in the value of buildings (B), their content (C) and profit they generated (P).

The data seem to be positively correlated, but it is not clear from the picture if there is a tail monotonic relation for a certain pair. As an example of a possible visual validation of the LTD property we present Figure 4.5. It depicts the slope of the C/u surface. The violation regions, where the slope is positive, are colored in orange and red. Clearly plots (a) and (b) show the largest violation regions. It is much harder to visually interpret the rest of the plots. What occurs from the testing procedure is that the tests reject neither $LTD(P|C)$ (c) nor $LTD(C|P)$ (f), but they reject $LTD(B|C)$ (d).

Table 4.4 contains approximate p-values for different pairs and tail configurations. Clearly none of the pairs has any considered negative dependence. For building and profit, and content and profit pairs none of the methods, with exception of the parametric Clayton reference method, rejects the null hypothesis of positive dependence in the right tails. As for the building and content pair only the parametric Frank method and the Π-reference method suggest positive dependence in the right tails. For this pair however, positive quadrant dependence was rejected by all of the testing methods considered in Chapter 3, and hence one expects that RTI should be rejected.

Note that both non-parametric methods lead to the same conclusions in this data example. This is in contrast to the parametric methods for which we see some huge discrepancies in the approximate p-values. This is in particular the case for $RTI(C|B)$, $RTI(B|C)$, $RTI(B|P)$, $LTD(P|C)$ and $LTD(C|P)$.

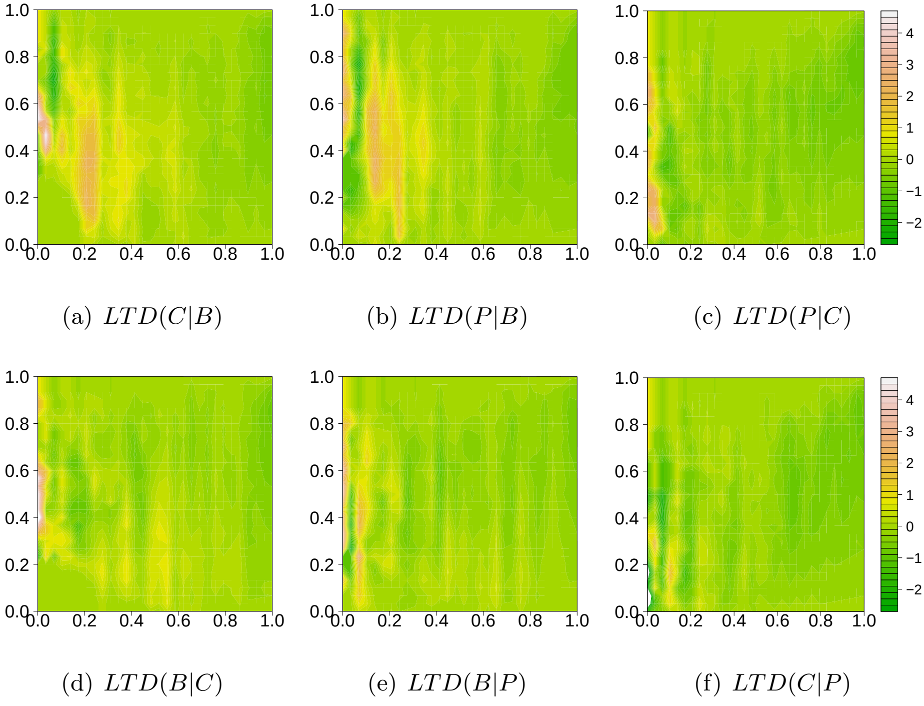


Figure 4.5: *Level plots of approximate u-slope of C/u surface.*

pair	method	$Y\|X$ LTD	$Y\|X$ RTD	$Y\|X$ LTI	$Y\|X$ RTI	$X\|Y$ LTD	$X\|Y$ RTD	$X\|Y$ LTI	$X\|Y$ RTI
	non-parametric I	0	0	0	0	0	0	0	0.002
(B,C)	non-parametric II	0	0	0	0	0	0	0	0.009
	parametric Frank	0	0	0.003	0.308	0	0	0	0.113
	parametric Clayton	0	0	0.003	0.030	0	0	0	0.006
	Π-reference	0	0	0.003	0.915	0.010	0	0	0.743
	non-parametric I	0	0	0	0.249	0.023	0	0	0.398
(B,P)	non-parametric II	0.001	0	0	0.230	0.063	0	0	0.432
	parametric Frank	0	0	0	0.802	0.005	0	0	0.557
	parametric Clayton	0	0	0	0.146	0	0	0	0.028
	Π-reference	0.002	0	0	1	0.543	0	0	0.998
	non-parametric I	0.383	0	0	0.510	0.767	0	0	0.987
(C,P)	non-parametric II	0.612	0	0	0.610	0.780	0	0	0.991
	parametric Frank	0.132	0	0	0.894	0.379	0	0	0.999
	parametric Clayton	0	0	0	0.262	0.004	0	0	0.882
	Π-reference	0.999	0	0	1	1	0	0	1

Table 4.4: *Danish fire insurance data:approximate p-values for all tail monotonicity tests.*

4.6.2 Market data

The second example is taken from the market data on the Belgian stock price index BEL20 (B), currency exchange rate EUR/DOL (ED) and gold index ETFS GOLD (G) in the period of two years from 30th June 2009 to 29th June 2011.

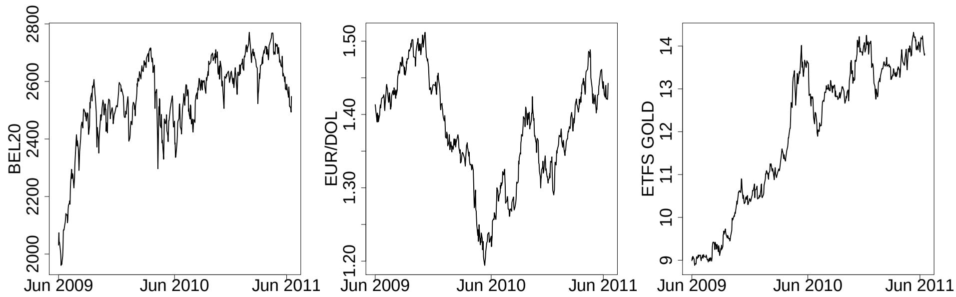
association measure	(B, ED)	(B, G)	(ED, G)
Pearson's cor.coef.	0.369	0.016	−0.234
Kendall's tau	0.244	0.022	−0.134
Spearman's rho	0.358	0.032	−0.195

Table 4.5: *Market data: empirical association measures.*

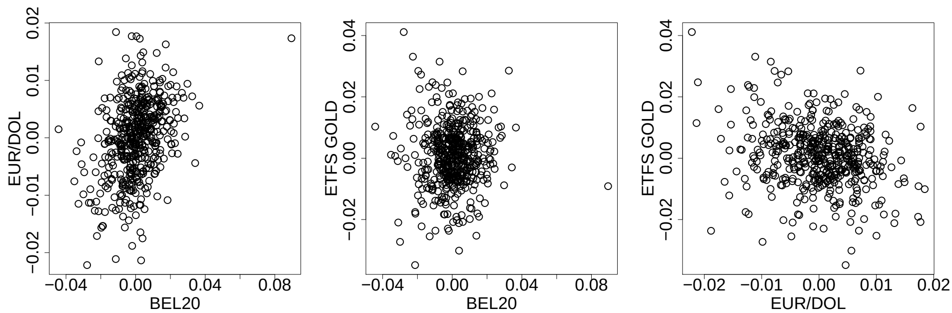
Figure 4.6 (a) contains plots of the original time series data. Parts (b) and (c) present log-returns (i.e., $\log\big(\mathrm{price}(t_{i+1})/\mathrm{price}(t_i)\big)$) and respective pseudo-log-returns in pairwise scatter plots. Tail monotonicity is in general hardly recognizable from the plots, yet we could expect from the BEL20 and EUR/DOL pair that it constitutes some positive dependence, whereas EUR/DOL and ETFS GOLD probably has some negative one. The remaining pair, BEL20 and ETFS GOLD, visually does not present any pattern. In Table 4.5 we summarize some global association measures for all pairs, and these suggest that there might be a global slightly positive dependence structure between BEL20 and ETFS GOLD.

The approximate p-values presented in Table 4.6 suggest that the (B, ED) pair is symmetrically positively tail dependent in both directions, whereas the (ED, G) pair is negatively dependent. For the pair (B, G) the results hardly reject any of the checked hypotheses. This might be an indicator that the sample is independent and this is indeed confirmed by external tests for independence.

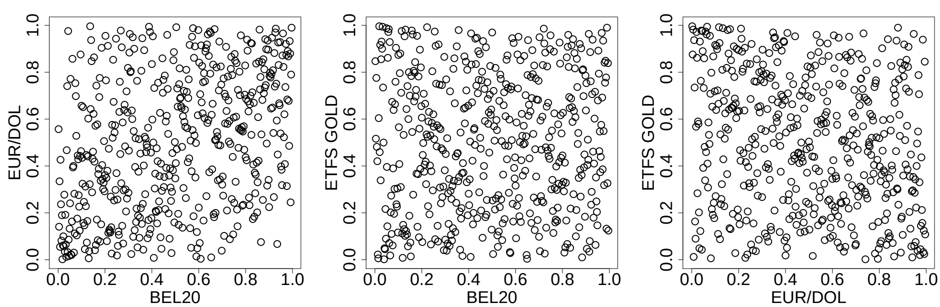
Note also for this example the large differences in approximate p-values for the two parametric methods, specifically for the cases $RTI(B|ED)$, $LTI(G|ED)$ and $LTI(ED|G)$.



(a)



(b)



(c)

Figure 4.6: *Market data. (a): original time series data; (b): log-returns pairwise; (c): pseudo-log-returns pairwise.*

		$Y\|X$				$X\|Y$			
pair	method	LTD	RTD	LTI	RTI	LTD	RTD	LTI	RTI
	non-parametric I	0.952	0	0	0.878	0.869	0	0	0.344
(B, ED)	non-parametric II	0.950	0	0	0.885	0.873	0	0	0.523
	parametric Frank	0.988	0	0	0.905	0.919	0	0	0.081
	parametric Clayton	0.464	0	0	0.161	0.287	0	0	0
	Π-reference	1	0	0	1	1	0	0	0.936
	non-parametric I	0.123	0.041	0.275	0.680	0.807	0.469	0.033	0.080
(B, G)	non-parametric II	0.265	0.105	0.334	0.771	0.871	0.490	0.056	0.138
	parametric Frank	0.169	0.066	0.408	0.858	0.944	0.739	0.049	0.062
	parametric Clayton	0.114	0.099	0.396	0.804	0.901	0.708	0.046	0.054
	Π-reference	0.328	0.084	0.404	0.943	0.981	0.726	0.045	0.165
	non-parametric I	0	0.735	0.462	0	0	0.891	0.416	0
(ED, G)	non-parametric II	0.001	0.796	0.622	0	0.001	0.917	0.521	0
	parametric Frank	0	0.771	0.281	0	0	0.953	0.201	0
	parametric Clayton	0	0.245	0.014	0	0	0.608	0.007	0
	Π-reference	0	0.998	0.928	0	0	1	0.891	0

Table 4.6: *Market data: approximate p-values for tail monotonicity tests.*

4.6.3 Air quality

The last data example concerns daily observations of four air quality measurements in New York from May until September 1973. The observations refer to mean Ozone parts (O), solar radiation (S), average wind speed (W) and maximum daily temperature (T). The data are analyzed in Chambers et al. (1983) and are publicly available in the *datasets* R package. The considered data set consists of 111 common observations.

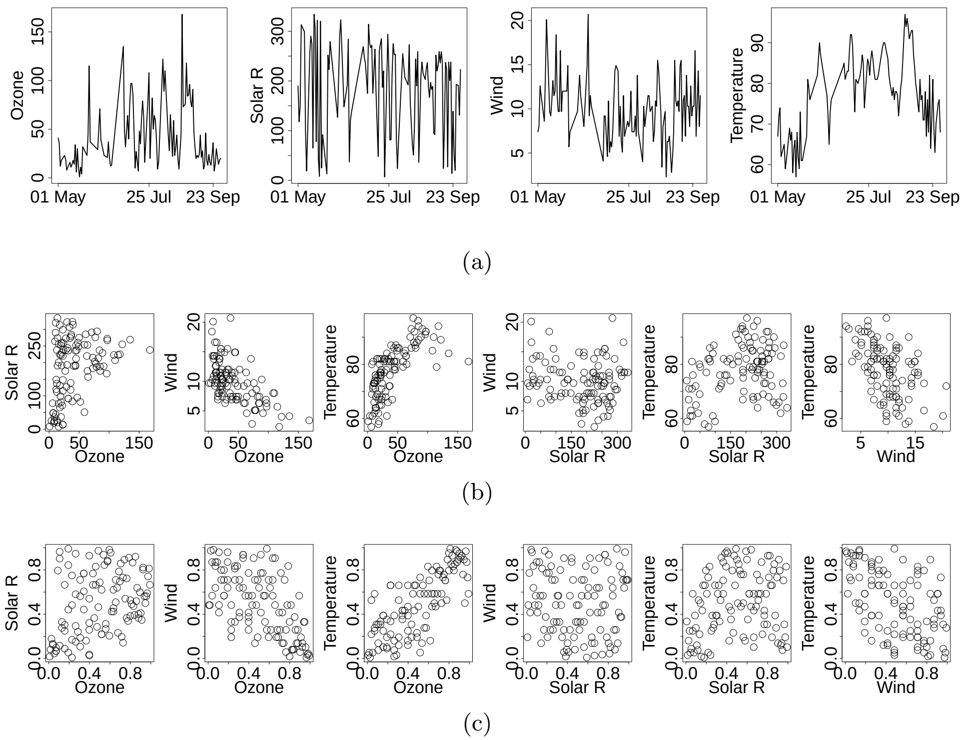


Figure 4.7: *Air quality data. (a): original time series data; (b): measurements pairwise; (c): pseudo-measurements pairwise.*

association measure	(O,S)	(O,W)	(O,T)	(S,W)	(S,T)	(W,T)
Pearson's cor.coef.	0.348	−0.612	0.698	−0.127	0.294	−0.497
Kendall's tau	0.240	−0.440	0.586	−0.043	0.142	−0.362
Spearman's rho	0.348	−0.605	0.772	−0.061	0.209	−0.499

Table 4.7: *Air quality data: empirical association measures.*

Graphical representations of the data in Figure 4.7 reveal different global dependence structures among the pairs. We can see some clear global positive relation in the (O, S) and (O, T) pairs and global negative ones in the (O, W) and (W, T) pairs, whereas the rest of the pairs exposes less visible global dependence structures. Table 4.7 summarizes some empirical association measures, and seems to confirm the global behaviour of pairs. However, Tables 4.8 and 4.9 reveal a lot more sophisticated dependence structures for some of the pairs. In particular, the $(W|S)$ distribution seems to be monotonic in both tails, yet with opposite impact, as we do not reject both $LTI(W|S)$ and $RTI(W|S)$. However, we do not reject any of the tests for $S|W$. The most sophisticated situation is with the (S, T) pair, where we do not reject $LTD(T|S)$ and $RTD(T|S)$, and nor $LTD(S|T)$ and $RTI(S|T)$.

There is only one significant discrepancy between the conclusions of the five testing methods. It is in testing for $RTI(O|S)$ where the parametric methods rejects the null hypothesis and the other methods do not (or barely not). As for the previous two examples, a similar remark can be made regarding different conclusions coming from the two parametric methods. See for example the cases of $RTD(O|W)$, $RTD(T|W)$, $LTD(T|O)$ and $LTD(O|T)$.

pair	method	$Y\|X$ LTD	RTD	LTI	RTI	$X\|Y$ LTD	RTD	LTI	RTI
(O,S)	non-parametric I	0.890	0	0	0.383	0.954	0.009	0	0.043
	non-parametric II	0.905	0.002	0	0.487	0.967	0.012	0	0.109
	parametric Frank	0.942	0.001	0	0.266	0.995	0.011	0	0.001
	parametric Clayton	0.749	0.002	0	0.057	0.955	0.014	0	0
	Π-reference	1	0.001	0	0.899	1	0.020	0	0.233
(O,W)	non-parametric I	0	0.965	0.711	0	0	0.698	0.977	0
	non-parametric II	0	0.959	0.763	0	0	0.764	0.984	0
	parametric Frank	0	0.976	0.469	0	0	0.266	0.999	0
	parametric Clayton	0	0.812	0.089	0	0	0.022	0.974	0
	Π-reference	0	1	1	0	0	0.989	1	0
(O,T)	non-parametric I	0.698	0	0	0.176	0.828	0	0	0.895
	non-parametric II	0.752	0	0	0.320	0.828	0	0	0.913
	parametric Frank	0.240	0	0	0.175	0.436	0	0	0.954
	parametric Clayton	0.018	0	0	0.010	0.051	0	0	0.741
	Π-reference	1	0	0	0.998	1	0	0	1

Table 4.8: *Air quality data: approximate p-values for tail monotonicity tests.*

		$Y\|X$				$X\|Y$			
pair	method	LTD	RTD	LTI	RTI	LTD	RTD	LTI	RTI
	non-parametric I	0.024	0.032	0.840	0.401	0.082	0.400	0.435	0.330
(S, W)	non-parametric II	0.032	0.068	0.866	0.409	0.125	0.511	0.467	0.327
	parametric Frank	0.035	0.019	0.964	0.688	0.109	0.425	0.575	0.376
	parametric Clayton	0.040	0.019	0.939	0.694	0.112	0.342	0.555	0.376
	Π-reference	0.016	0.040	0.970	0.781	0.154	0.582	0.786	0.323
	non-parametric I	0.847	0.127	0	0.019	0.484	0.005	0.002	0.525
(S, T)	non-parametric II	0.877	0.119	0	0.060	0.539	0.016	0.004	0.591
	parametric Frank	0.984	0.229	0	0.002	0.591	0.013	0.003	0.383
	parametric Clayton	0.939	0.211	0	0	0.352	0.018	0	0.168
	Π-reference	1	0.174	0	0.070	0.921	0.015	0.001	0.835
	non-parametric I	0	0.587	0.842	0	0	0.695	0.668	0
(W, T)	non-parametric II	0	0.675	0.844	0	0	0.721	0.682	0
	parametric Frank	0	0.153	0.957	0	0	0.798	0.546	0
	parametric Clayton	0	0.009	0.719	0	0	0.367	0.147	0
	Π-reference	0	0.974	1	0	0	1	0.997	0

Table 4.9: *Air quality data: Table 4.8 continued.*

4.7 Conclusions and further discussion

In this chapter we introduced testing procedures for testing for a specific tail monotonicity dependence structure. We propose several finite-sample resampling approaches for obtaining approximate p-values for the test statistic. A simulation study reveals that the Π-reference method might occasionally give very inaccurate results. The performance of the parametric reference resampling approach can depend very much on the choice of the parametric reference family as was seen from the simulation study as well as from the data examples. On the one hand, this sensitivity to the specification of the parametric copula reference family is a clear drawback of this parametric resampling approach. On the other hand, the computational advantage is a plus in comparison with the non-parametric method. We presented several real data examples, where the various tests often reveal several interesting and sophisticated compositions of dependence structures among the variables.

In the next chapter, we present a constrained copula approach to testing for the strongest among the specific dependence structures discussed in the introduction, namely stochastic monotonicity.

Chapter 5

Testing stochastic monotonicity by constrained copula estimation

5.1 Introduction

This chapter is based on Gijbels and Sznajder (2011c) and develops tests for stochastic monotonicity. Recall, that Y is stochastically increasing in X when $\mathbb{P}(Y \leq y | \; X = x)$ is a non-increasing function in x for every y.

The stochastic monotonicity concept originates from the work of Tukey (1958) and Lehmann (1959), where it was called complete regression dependence. This name refers to the idea behind the concept, namely that all conditional quantiles are monotonic functions (in the same direction). In particular it implies what is nowadays called regression monotonicity, i.e., monotonicity of $\mathbb{E}(Y | \; X = x)$ as a function of x. This property has been studied by many authors, e.g., Bowman et al. (1998), Gijbels et al. (2000), Birke and Dette (2007) and Antoniadis et al. (2007).

The concept of stochastic monotonicity has gained particular interest in econometrics, where it is assumed in many studies. This eventually yields a question on justification of the assumption of stochastic monotonicity. A recent answer to this question was given in Lee et al. (2009), where the authors construct a statistical test for testing for stochastic increasingness. This article also includes a broad overview of recent applications of the stochastic monotonicity concept in econometrics.

The method in Lee et al. (2009) for testing for stochastic monotonicity is supported by an asymptotic distribution of a supremum of the following rescaled U-statistic

$$\begin{aligned}\widehat{U}_n(x,y) = &\frac{2}{n(n-1)} \sum_{1\le i<j\le n} \big[\mathbb{I}(Y_i \le y) - \mathbb{I}(Y_j \le y)\big]\operatorname{sgn}(\widehat{X}_i - \widehat{X}_j) \\ &\cdot k_{h_n}(\widehat{X}_i - x)k_{h_n}(\widehat{X}_j - x),\end{aligned}$$

where $\widehat{X}_i = \Psi(W_i, \widehat{\theta})$ is coming from the assumption of the authors that in the sample $\big\{(X_i, Y_i)\big\}_{i=1}^n$ the first argument is unobservable, yet it is a known parametric function Ψ of an observable variable W, and $\widehat{\theta}$ is an n-root consistent estimator of the true parameter θ. The signum function is defined as $\operatorname{sgn}(x) = \mathbb{I}(x \ge 0) - \mathbb{I}(x \le 0)$ and the rescaled kernel function as $k_{h_n}(x) = \frac{1}{h_n} k\left(\frac{x}{h_n}\right)$ for a given kernel function k and bandwidths $h_n > 0$. This U-statistic can, in the authors' words, "be viewed as a locally weighted version of Kendall's tau statistic applied to $\mathbb{I}(Y \le y)$". Furthermore, it can be seen as a discrete (rescaled) approximation of $F_x(y|x)$, the partial derivative of $\mathbb{P}(Y \le y|\ X = x)$ with respect to x. Now, if we look at the definition of stochastic increasingness in Definition 1.14, we can see that (under some smoothness conditions) this definition is equivalent to $F_x(y|x)$ being non-positive for every x. Therefore if the supremum of the U-statistic is "too positive" it suggests rejection of the null hypothesis of stochastic increasingness in the data.

As we have seen in Chapters 2 and 3, the discrete partial derivative estimation as an approximation to the asymptotic expression can be sometimes lacking power in small to moderate sample size problems. In this chapter we thus incorporate the non-parametric resampling methodology developed in Chapters 3 and 4 in the stochastic monotonicity testing problem, i.e., in testing

$$\begin{aligned}&H_0 : \ Y \text{ is stochastically increasing in } X \\ &\text{versus} \\ &H_1 : \ Y \text{ is not stochastically increasing in } X.\end{aligned} \tag{5.1}$$

Stochastic monotonicity as a feature of dependence structure is a strengthening of the tail monotonicity concept (see Proposition 1.14). In other words, when not rejecting the null hypothesis of left tail decreasingness, we might get further insight in the data, by testing for stochastic increasingness.

This chapter is organized as follows. Section 5.2 contains the theoretical definition and properties of the stochastic monotonic dependence structure. In Section 5.3 we motivate the construction of the test statistic. Section 5.4

describes the constrained SI adjustment and constrained resamling technique. Simulation study results are discussed in Section 5.5 and Section 5.6 investigates two data examples. Section 5.7 concludes the chapter with a discussion.

5.2 Stochastic monotonicity

Stochastic monotonicity is the strongest dependence feature studied in this thesis. It is defined as stochastic increasingness in Definition 1.14, but in a similar way one can define a stochastic decreasing relation. Stochastic monotonicity is not a symmetric concept with respect to the variables. As we have seen in Proposition 1.15, the relation that a random variable Y is stochastically increasing in another random variable X, can be expressed as a feature of the underlying copula function; namely, that it is a concave function in the first argument for every fixed value of the second argument.

In the next section we will focus on developing a statistical test for (5.1). However, as a consequence of Proposition 5.1, we can apply the same test to a monotonically transformed data set to test also for stochastic decreasingness. Similarly to Proposition 4.1 one can prove the following statements.

Proposition 5.1. *Let α be a strictly increasing real function, and β_1 and β_2 be strictly decreasing real functions. Then*

(a) $\alpha(Y)|\beta_1(X)$ is stochastically increasing if and only if $Y|X$ is stochastically decreasing, i.e.,

$$SI\left(\alpha(Y)|\beta_1(X)\right) \Longleftrightarrow SD(Y|X)$$

(b) $\beta_2(Y)|\alpha(X)$ is stochastically increasing if and only if $Y|X$ is stochastically decreasing, i.e.,

$$SI\left(\beta_2(Y)|\alpha(X)\right) \Longleftrightarrow SD(Y|X)$$

(c) $\beta_2(Y)|\beta_1(X)$ is stochastically increasing if and only if $Y|X$ is stochastically increasing, i.e.,

$$SI\left(\beta_2(Y)|\beta_1(X)\right) \Longleftrightarrow SI(Y|X).$$

Finally, note that we can express the condition for SI in a different way than in Proposition 1.15.

Proposition 5.2. *$SI(Y|X)$ if and only if for every v, $\frac{\partial C(u,v)}{\partial u}$ is a non-increasing function of u.*

Proposition 5.2 deals with a condition on a partial derivative of a function and this is much harder to impose and/or estimate than a condition on the function itself.

5.3 Test statistic

From Proposition 1.15 we have that problem (5.1) is equivalent to

$H_0:\ C(u,v)$ is a concave function in u for every v
versus
$H_1:\ C(u,v)$ is not a concave function in u for some v.

The test statistic that we propose here is motivated by theoretical considerations of a certain distance between a copula and its concave (only in the u direction) hull. Specifically, let us define $\breve{\cdot}$ to be an operator on a set of bounded functions, which transforms a function $f:\ A \to B$ to its smallest concave majorant, i.e.,

$$\breve{f}(x) = \inf\big\{g(x)|\ g:\ A \to B\ \&\ g \text{ is concave on } A\big\}. \tag{5.2}$$

Using (5.2) we specify the following distance to be the measure of discrepancy in the SI testing problem

$$\int \big(\breve{C}(\cdot,v)(u) - C(u,v)\big)dudv. \tag{5.3}$$

Naturally, in the testing problem none of the functions in (5.3) is known, thus we consider their empirical versions and get to the test statistic

$$T_n = \int \big(\breve{C}_n(\cdot,v)(u) - C_n(u,v)\big)dudv, \tag{5.4}$$

where C_n is the empirical copula estimator defined in (1.3).

To obtain critical values for the given test statistic we propose to apply the method of constrained copula smoothing and resampling, described in detail in Section 5.4. The basic idea behind the proposed inference, is to draw ($j = 1, \ldots, N$) samples (of size n) from a constrained copula estimator and for each compute the corresponding test statistic $T^*_{n,j}$, and approximate the p-value of the original test statistic as

$$p^* = \frac{1}{N}\sum_{j=1}^{N} \mathbb{I}\big(T^*_{n,j} > T_n\big). \tag{5.5}$$

5.4 SI adjustment and resampling

We will apply the ideas of Chapters 3 and 4 and focus on the function $\breve{C}$ from (5.3). Note that if for a fixed v, C is a concave function of u, then $\breve{C}(\cdot, v) \equiv C(\cdot, v)$. In particular, if all of the v-sections of C are concave then $\breve{C} \equiv C$. In general however, $\breve{C}$ might not be a valid distribution function. Therefore, we apply the rearrangement technique described in Section 1.2.3 to the partial derivative of $\breve{C}$ and integrate it back to obtain a true distribution function, and finally, a copula.

5.4.1 SI adjustment

Figure 5.1 depicts a scheme of the steps that lead to the construction of a constrained copula.

$$
\begin{tikzcd}
C(u,v) \arrow[ddd, "1"'] & & & & \\
& \breve{c}_u(v) \arrow[ddd, "3"'] & & & \\
& & \widetilde{C}(u,v) \arrow[ddd, "5"'] \arrow[drr, "7"] & & \\
\breve{C}(u,v) \arrow[uur, "2"] & & & & \breve{\widetilde{C}}(u,v) \\
& \widetilde{c}_u(v) \arrow[uur, "4"] & & & \\
& & G_{\widetilde{V}}(v) \arrow[uurr, "6"] & &
\end{tikzcd}
$$

Figure 5.1: *Schema of the construction of a constrained copula.*

1. Having a copula C we apply the $\breve{\cdot}$ operator to obtain the "concave hull" $\breve{C}$ of C.

2. Then, we calculate the partial derivative of $\breve{C}$ with respect to u

$$\breve{c}_u(v) = \frac{\partial \breve{C}(u,v)}{\partial u}$$

3. and monotonize it so that for every u it is a proper univariate distribution function on the unit interval

$$\widetilde{c}_u(v) = (\xi \circ \xi)\big(\breve{c}_u(\cdot)\big)(v). \tag{5.6}$$

4. Then, we integrate $\widetilde{c}_u(v)$ to obtain a valid distribution function on the unit square $(U, \widetilde{V}) \sim \widetilde{C}$.

$$\widetilde{C}(u, v) = \int_0^u \widetilde{c}_t(v) dt. \tag{5.7}$$

The marginal distribution of the first component of $\widetilde{C}$ is the uniform distribution on the unit interval, by construction,

5. and the marginal distribution of the second component can be computed as follows

$$G_{\widetilde{V}}(v) = \widetilde{C}(1, v).$$

6,7. Eventually, we can write the copula function of the distribution $\widetilde{C}$ as

$$\breve{\widetilde{C}}(u, v) = \widetilde{C}\big(u, G_{\widetilde{V}}^{-1}(v)\big).$$

Note that if the second order partial derivative $\frac{\partial^2 C(u,v)}{\partial u^2}$ of the copula C exists, then it has to be non-positive for the concavity constraint to hold, see also Proposition 5.2. In Figure 5.2 we present contour plots of the second order partial derivative (with respect to u) of the copula B from Section 4.5 and the (approximate) second order partial derivative of the corresponding SI-constrained copula $\breve{\widetilde{C}}$.

From Figure 5.2(b) we can see that imposing the constraint works, as there are no positive levels anymore. However, the general shape of the derivative visibly changed and thus in this case $\breve{C}$ and $\breve{\widetilde{C}}$ are significantly different.

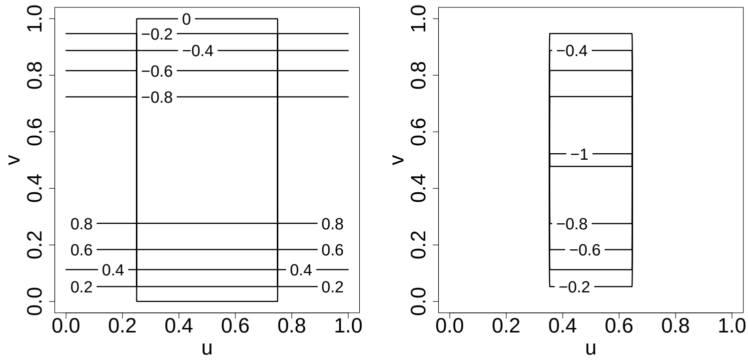
5.4.2 Constrained resampling

The idea behind constrained resampling is to estimate $\widetilde{c}_u(v)$ in (5.6) and follow Algorithm 1.1 to obtain samples from $(U, \widetilde{V}) \sim \widetilde{C}$ in (5.7) as

$$\widetilde{c}_u(v) = \mathbb{P}(\widetilde{V} \leq v |\, U = u).$$

We first estimate $\breve{c}_u(v)$ by means of a local polynomial method as discussed in Section 1.2.3. More precisely, for a fixed grid of points in the unit square $\big\{(u_i, v_j)\big\}_{i,j=0}^{m+1}$ we take

$$c_{i,j} = \breve{C}_n(u_i, v_j) \tag{5.8}$$



(a) (b)

Figure 5.2: *Contour plots of the second order partial derivative* $\frac{\partial^2 C}{\partial u^2}$ *of copula B from Section 4.5 (a) and of the corresponding constrained copula* $\breve{\tilde{C}}$ *(b).*

and then apply the rearrangement operator to the obtained values $\breve{c}_{n,u}(v)$ to get an estimator $\widetilde{c}_{u,n}(v)$ of (5.6)

$$\widetilde{c}_{u,n}(v) = (\xi \circ \xi)\big(\breve{c}_{n,u}(\cdot)\big)(v).$$

Due to the finite sample error this estimate is likely not to satisfy the boundary conditions of a univariate distribution function on the unit interval as opposed to its theoretical counter-part (5.6), thus we transform it linearly to fix the end point values in zero and one, i.e.,

$$\widehat{\widetilde{c}}_{u,n}(v) = \frac{\widetilde{c}_{u,n}(v) - \widetilde{c}_{u,n}(0)}{\widetilde{c}_{u,n}(1) - \widetilde{c}_{u,n}(0)}.$$

Eventually, $\widehat{\widetilde{c}}_{u,n}(v)$ is a proper (random) univariate distribution function and is used for the resampling process and the calculation of the approximate p-value as in (5.5).

In the next section we present part of a large simulation study that was conducted to verify the power and size performance of the described testing procedure and compare it with the proposed method of Lee et al. (2009). Similarly as in Chapter 4 we shall see that the non-parametric procedure proposed so far creates resampled samples which are "too SI". This results generally in a greater chance of rejection, even under the null hypothesis. Specifically, the simulation results for the independence copula are significantly above the prescribed level of the test.

Therefore, we propose an alternative to obtain values of $c_{i,j}$ in (5.8) as a starting point of the procedure. The algorithm of Meyer (2008) described in Section 4.4 allows for a constrained least squares fit of a general problem

$$\hat{\zeta} = \text{argmin}||\mathbf{x} - \zeta||^2 \quad \text{such that} \quad A\zeta \geq 0, \tag{5.9}$$

where $\mathbf{x}$ is a vector of observations and A is an irreducible matrix of linear conditions.

We apply the cone projection algorithm of Meyer (2008) to the points

$$\mathbf{x} = \{x_i\}_{i=0}^{m+1} = \{C_n(u_i, v_j) - u_i v_j\}_{i=0}^{m+1}$$

for every fixed j. Note that we want $\mathbf{x}$ to form a concave shape grounded to zero in the end points, for the boundary condition to hold, i.e., $C_n(1, v_j) = v_j$. As for the concavity, we require that for $i = 0, \ldots, m-1$

$$x_{i+2} - x_{i+1} \leq x_{i+1} - x_i. \tag{5.10}$$

The two constraints yield the following condition matrix A

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & \ldots & 0 & 0 & 0 & 0 \\ & & \ddots & & & \ddots & & \ddots & \\ 0 & 0 & 0 & 0 & \ldots & 0 & -1 & 2 & -1 \end{bmatrix},$$

where the first two rows correspond to the condition related to the boundary values and the other rows to the concavity condition. The copula estimator that is obtained with the constrained least squares fit is denoted by C_n^{LSC}.

Figure 5.3 depicts differences between the estimators from the copula B in Section 4.5. We can see that C_n^{LSC} is "less SI", than the concave envelope $\breve{C}_n$.

Note that one can also follow a parametric resampling approach similar to these discussed in the previous chapters. The Frank and Clayton copula families are again SI for parameter values not smaller than zero and SD for parameter values not greater than zero. To sum up, we shall include in the simulation study the following testing procedures:

NON-PARAMETRIC I: based on $\breve{C}_n$ in (5.8);

NON-PARAMETRIC II: based on C_n^{LSC} from the cone projection algorithm;

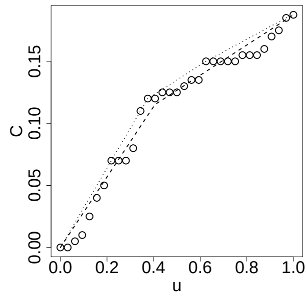


Figure 5.3: *Illustration of copula adjustment for the copula B in Section 4.5; section of the empirical copula C_n for fixed v (points), of the corresponding C_n^{LSC} (dashed line) and of the $\breve{C}_n$ (dotted line) values.*

PARAMETRIC FRANK: using parametric resampling from Frank copula family;

PARAMETRIC CLAYTON: using parametric resampling from Clayton copula family;

Π-REFERENCE: using the independence copula for the resampling;

LEE'S METHOD: method described in Lee et al. (2009).

5.5 Simulation study

In this section we present results of the simulation study on power and size performance of the proposed testing procedures.

We consider several distribution examples to generate 1000 samples of size $n = 200$ and $n = 400$. For each sample another 1000 iterations were used for the resampling and for assessing the p-values. The significance level was set to 0.05 and we report on the percentage of rejected tests.

In the non-parametric methods we used the grid size of $m = \lfloor \sqrt{n} \rfloor$ for approximating the integral in the test statistic and for estimating the smooth partial derivative. The bandwidth parameters were taken equal to $1.5/m$.

We present results for the Frank(-1) copula, the independent pair of random variables, the copula B introduced in Section 4.5 and three distributions constructed by specifying the mean regression function, i.e.,

(S1) a distribution example considered in Lee et al. (2009)

$$X \sim U[0;1] \quad Y|X = \mathcal{N}\big(X(1-X), .1\big)$$

(S2)

$$X \sim U[-3; 3] \quad Y|X \sim \mathcal{N}\big(\sin X, 1\big)$$

(S3)

$$X \sim U[0; 1] \quad Y|X \sim \mathcal{N}\big(0.3 \sin(2\pi X) + X, X(1 - X) + 0.1\big).$$

In Figure 5.4 we present exemplary samples from copulas $S1$, $S2$ and $S3$ together with the corresponding pseudo-observations. Naturally, from the construction of these copula distributions we have that for none of them Y is stochastically increasing or decreasing in X, as the mean regression function is not a non-decreasing function. A summary of the stochastic monotonicity structures that we investigate in this simulation study is given in Table 5.1.

	$Y\|X$		$X\|Y$	
copula	SI	SD	SI	SD
Frank(−1)		×		×
Π-copula	×	×	×	×
copula B			×	
S1				
S2			×	
S3			×	

Table 5.1: *The copula simulation models and their dependence structures.*

Table 5.2 contains the results of the simulation study for sample size 200.

For the "classical" Frank copula example, we can see that all of the proposed methods reveal similar powers. The percentage of rejections when the hypothesis is true, in case of $SD(Y|X)$ or $SD(X|Y)$, is always below the prescribed significance level of the tests. Lee's method reveals less power. This might be caused by the fact that this testing procedure is rather "conservative" judging on the simulation study results for the independence copula, i.e., it rejects fewer samples than the prescribed five percent level.

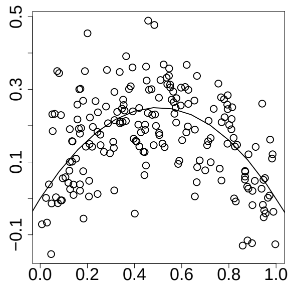
As was mentioned in Section 5.4, the non-parametric method I does not hold the level. The rest of the methods seems to oscillate around the five percent level in the independence copula example.

For the copula B we can see that it is much harder to detect $SI(Y|X)$ than $SD(Y|X)$. Especially, we can observe that the Π-reference method and Lee's method do not seem to work at all. For the $X|Y$ dependence all of the methods perform equally well.

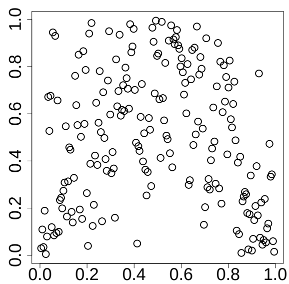
For the copula $S1$ all the approaches correctly reject a vast majority of samples and we can see that Lee's method outperforms the Π-reference method, yet its power is below the method proposed in this chapter.

The results for the Copulas $S2$ and $S3$ are very similar to those for copula B, with the exception of the results for testing for the $SI(X|Y)$ in Coupla $S3$. The same situation is seen for Lee's method for copula $S3$.

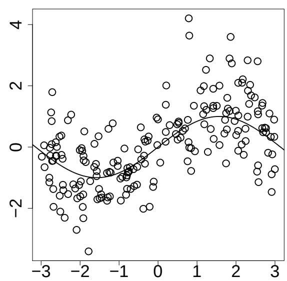
The results when moving to a larger sample size ($n = 400$) are presented in Table 5.3. The evolution in power is visible for all the methods in almost all of the examples. The exception is the Π-reference method, which does not seem to improve at all for $SI(Y|X)$ and Copulas B, $S2$ and $S3$. The same situation is with Lee's method for copula $S3$. Note also the slow convergence to the true distribution for Lee's method, seen from the results for the independence copula, as they are again much lower than the five percent significance level.



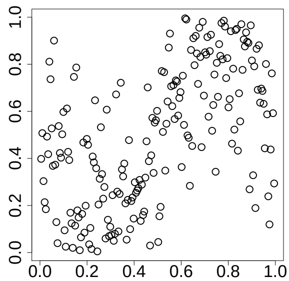
(a)



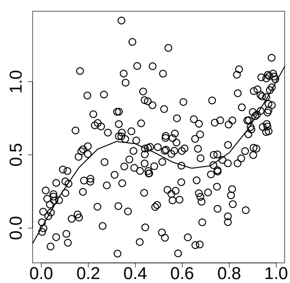
(b)



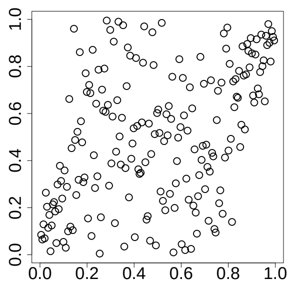
(c)



(d)



(e)



(f)

Figure 5.4: *Samples (left column) of size* 200 *from Copulas* S_1, S_2 *and* S_3 *(top, middle and bottom row accordingly) with indicated lines of mean conditional functions, and the corresponding pseudo-observations (right column).*

copula	method	$Y\|X$ SI	$Y\|X$ SD	$X\|Y$ SI	$X\|Y$ SD
Frank(−1)	non-parametric I	0.751	0.011	0.766	0.020
	non-parametric II	0.690	0.002	0.708	0.001
	parametric Frank	0.701	0.011	0.707	0.011
	parametric Clayton	0.705	0.003	0.713	0.006
	Π-reference	0.702	0	0.713	0
	Lee's method	0.315	0.001	0.316	0.002
Independence	non-parametric I	0.117	0.129	0.126	0.116
	non-parametric II	0.049	0.052	0.055	0.059
	parametric Frank	0.052	0.047	0.059	0.055
	parametric Clayton	0.051	0.046	0.054	0.055
	Π-reference	0.052	0.049	0.053	0.058
	Lee's method	0.031	0.029	0.028	0.033
B	non-parametric I	0.354	0.999	0.030	0.984
	non-parametric II	0.200	0.996	0.002	0.981
	parametric Frank	0.431	0.996	0.036	0.984
	parametric Clayton	0.296	0.996	0.012	0.986
	Π-reference	0	0.996	0	0.986
	Lee's method	0.052	0.967	0	0.908
$S1$	non-parametric I	1	1	0.970	0.971
	non-parametric II	1	1	0.939	0.954
	parametric Frank	1	1	0.939	0.957
	parametric Clayton	1	1	0.928	0.940
	Π-reference	0.995	0.997	0.619	0.643
	Lee's method	1	1	0.728	0.749
$S2$	non-parametric I	0.674	1	0.023	1
	non-parametric II	0.436	1	0.008	1
	parametric Frank	0.925	1	0.052	1
	parametric Clayton	0.869	1	0.019	1
	Π-reference	0	1	0	1
	Lee's method	0.069	1	0	1
$S3$	non-parametric I	0.750	1	0.006	1
	non-parametric II	0.499	1	0	1
	parametric Frank	0.957	1	0.055	1
	parametric Clayton	0.934	1	0.022	1
	Π-reference	0.006	1	0	1
	Lee's method	0	1	0.143	0.995

Table 5.2: *Simulation study results for sample size* 200.

		$Y\|X$		$X\|Y$	
copula	method	SI	SD	SI	SD
Frank(1)	non-parametric I	0.956	0	0.940	0.014
	non-parametric II	0.925	0	0.933	0.008
	parametric Frank	0.958	0	0.935	0.022
	parametric Clayton	0.901	0	0.929	0.004
	Π-reference	0.928	0	0.939	0
	Lee's method	0.604	0	0.658	0
Independence	non-parametric I	0.110	0.105	0.106	0.081
	non-parametric II	0.055	0.020	0.057	0.022
	parametric Frank	0.052	0.039	0.055	0.019
	parametric Clayton	0.034	0.020	0.062	0.012
	Π-reference	0.046	0.018	0.053	0.020
	Lee's method	0.005	0.032	0.026	0.008
B	non-parametric I	0.673	1	0.027	1
	non-parametric II	0.471	1	0	1
	parametric Frank	0.805	1	0.054	1
	parametric Clayton	0.690	1	0.020	1
	Π-reference	0	1	0	1
	Lee's method	0.112	1	0.002	1
$S1$	non-parametric I	1	1	1	1
	non-parametric II	1	1	1	1
	parametric Frank	1	1	1	1
	parametric Clayton	1	1	1	1
	Π-reference	1	1	0.961	0.976
	Lee's method	1	1	0.990	0.982
$S2$	non-parametric I	0.957	1	0	1
	non-parametric II	0.863	1	0	1
	parametric Frank	0.999	1	0.041	1
	parametric Clayton	1	1	0	1
	Π-reference	0.002	1	0	1
	Lee's method	0.121	1	0.002	1
$S3$	non-parametric I	0.943	1	0.003	1
	non-parametric II	0.788	1	0.003	1
	parametric Frank	0.990	1	0.072	1
	parametric Clayton	0.994	1	0.013	1
	Π-reference	0.011	1	0.005	1
	Lee's method	0	1	0.347	1

Table 5.3: *Simulation study results for sample size* 400.

5.6 Real data examples

In this section we illustrate the application of the described tests to real data examples.

5.6.1 Danish fire insurance data

The Danish fire insurance data example introduced in Section 3.5 consists of observations on three random variables referring to insurance losses in the value of buildings (B), their content (C) and profit they generated (P). As seen in Table 4.4 we rejected any tail dependence for the couple $\{B, C\}$, we have seen a positive right tail dependence for the couple $\{B, P\}$, and we have decided upon a positive dependence in both tails for the couple $\{C, P\}$. Thus, referring to Proposition 1.14 it is only worth to investigate further the $\{B, P\}$, and $\{C, P\}$ pairs for stochastic monotonicity. Indeed, we do not include in Table 5.5 approximate p-values for the couple $\{B, C\}$ as all methods strongly reject the null hypotheses of any stochastic monotonic dependence (returning approximate p-values equal to zero). The same holds for any negative stochastic monotonicity structure for all the couples, and so we drop these zero-columns from Table 5.5. Note that we wish to obtain consistent results with respect to Proposition 1.14. In particular we would like (potentially) not to reject only cases of SI for which we have not rejected both LTD and RTI.

The more interesting findings concern the other two pairs $\{B, P\}$ and $\{C, P\}$. The hypothesis of $SI(P|B)$ is rejected by all the methods (except the Π-reference one), which is as expected as all the methods strongly rejected $LTD(P|B)$. The results are also consistent for $SI(B|P)$: the non-parametric II and Π-reference methods did not reject both $LTD(B|P)$ and $RTI(B|P)$ and only they do not reject $SI(B|P)$.

The $\{C, P\}$ pair seems to have more symmetric dependence structure as none of the tests reject the stochastic increasingness neither for $P|C$ nor for $C|P$. However, there is one inconsistency in the results for this couple. Specifically, the parametric Clayton method strongly rejects $LTD(P|C)$ and $LTD(C|P)$, but it rejects neither $SI(P|C)$ nor $SI(C|P)$. This situation is another example of the importance of correctly specifying the model in the parametric resampling method.

Similar plots to the ones in Figure 4.5 were drawn in case of SI, where the violation regions were referring to positive approximated second order partial derivative of C with respect to u. They were however unreadable and therefore uninformative.

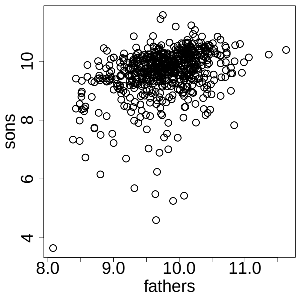
method	$(P\|B)$	$(B\|P)$	$(P\|C)$	$(C\|P)$
non-parametric I	0.0001	0.0421	0.9086	0.4972
non-parametric II	0.0020	0.0786	0.9465	0.6486
parametric Frank	0	0.0013	0.7759	0.0804
parametric Clayton	0	0.0095	0.8504	0.1718
Π-reference	0.1603	0.9117	1	1

Table 5.4: *Approximate p-values for the SI tests for the Danish Fire Insurance data example.*

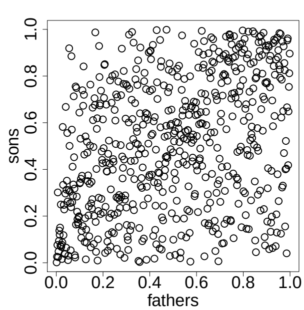
5.6.2 Intergenerational income data

This data set was tested for stochastic increasingness in Lee et al. (2009). The data come from the Panel Study of Income Dynamics (PSID), first introduced by Solon (1992). We analyze the same data subsample as in Lee et al. (2009), which is a data extract from Minicozzi (2003) (available online on the *Journal of Applied Econometrics* website).

The data consist of 616 observations of 'the logarithm of parental predicted permanent income' (variable X) and 'the logarithm of sons averaged full-time real labor income at ages 28 and 29' (variable Y). The scatter plot of the data and pseudo-observations is presented in Figure 5.5. We can clearly see some positive dependence in the data and indeed there is a very strong positive dependence structure as the tests do not reject any of the methods, see Table 5.5. This is in line with the conclusions in Lee et al. (2009).



(a)



(b)

Figure 5.5: *Income data (a) and corresponding pseudo-observations (b).*

method	$(Y\|X)$	$(X\|Y)$
non-parametric I	0.8705	0.8762
non-parametric II	0.9309	0.9411
parametric Frank	0.6191	0.5409
parametric Clayton	0.7491	0.6812
Π-reference	1	1

Table 5.5: *Approximate p-values for the SI test for the Income data example.*

5.7 Conclusions and further discussion

In this chapter we introduced a non-parametric test for a stochastic monotonicity dependence structure based on the constrained copula estimation and resampling.

In a simulation study we evaluated the power performances of the discussed testing procedures, and compared these with the performance of the method of Lee et al. (2009). The latter method relies on the asymptotic distribution of a U-statistic type of testing quantity.

We applied the tests to the real data examples. We have obtained interesting dependence structure results for the Danish fire insurance data which gives more insight into the data set. We have also found evidence in the available data for the assumption of the stochastic increasingness in the econometric model of intergenerational income.

The conclusions from both the simulation studies and the real data examples are that the Π-reference method is not a valid approach any more. The class of stochastically monotonic structures is “too narrow” to be approximated by the Π copula boundary point. Finally, we advise to use the non-parametric method II if there is no evidence or a prior knowledge of the possible parametric copula model as it might also sometimes lead to inaccurate results.

Chapter 6

General conclusions and perspectives

This thesis describes the research done on dependence structure testing problems. It explores the copula applicability and constructs non-parametric tests by means of smooth constrained resampling.

In Chapter 2 we start the analysis of the weakest form of dependence, namely Positive Quadrant Dependence. Different copula estimators are used to construct the test statistic and the rejection rule is based on resampling from the Π-reference copula. The Π-reference method improves upon the existing method of Scaillet (2005) and kernel copula estimators combined with Cramér-von Mises and Anderson-Darling distances outperform tests based on the empirical copula estimator and Kolmogorov-Smirnov distances.

In Chapter 3 we consider a constrained copula resampling approach in Positive Quadrant Dependence testing. We manage to improve the power performance further and introduce parametric resampling for comparison.

Chapter 4 applies the constrained copula resampling methodology to another dependence structure, namely Tail Monotonicity. This dependence feature has not yet been, up to the authors' knowledge, an object in statistical testing. The advantage of the non-parametric approach over the Π-reference and parametric approaches is more visible when investigating this stronger dependence structure of Tail Monotonicity. On the one hand, the Π-reference method does not seem to catch the details which constitute the essence of this dependence structure. On the other hand, the test is sensitive to misspecification of the parametric reference copula model used for resampling.

These drawbacks emerge even further in the case of testing for Stochastic

Monotonicity in Chapter 5. Moreover, the proposed non-parametric method outperforms the recently developed testing method of Lee et al. (2009).

All of the dependence structures are tested on a variety of simulation examples with broad ranges of difficulties to give an idea about applicability of the tests. In the real data examples the tests demonstrate clearly their use in broadening the knowledge about the association in these data sets.

A challenging topic for further research is to prove in general the methodology summarized in Figure 5.1. It would give a tool to handle other possible dependence structures. The main difficulty here is the impact of the rearrangement step on the theoretical features of the transformed copula. With such a result at hand one can prove statistical consistency of the methods.

Another important question is the theoretically based bandwidth selection procedure for testing. It is an important and unexplored topic in general and in particular in copula studies. The large amount of conducted simulations suggests however strong undersmoothing, which is a direction for the investigation in the area of the constrained resampling for testing.

With the above questions answered, one receives a convenient and coherent tool to be applied to testing for any dependence structures. For example, one can consider local behaviour of the described dependence structures as very often global features are too restrictive and certain regions in the data range are of no/most interest.

Bibliography

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, 49:765–769.

Antoniadis, A., Bigot, J., and Gijbels, I. (2007). Penalized wavelet monotone regression. *Statistics & Probability Letters*, 77(16):1608–1621.

Berg, D. (2009). Copula goodness-of-fit testing: an overview and power comparison. *European Journal of Finance*, 15(7–8):675–701.

Birke, M. and Dette, H. (2007). Testing strict monotonicity in nonparametric regression. *Mathematical Methods of Statistics*, 16(2):110–123.

Bowman, A. W., Jones, M. C., and Gijbels, I. (1998). Testing monotonicity of regression. *Journal of Computational and Graphical Statistics*, 7(4):489–500.

Braeken, J. and Tuerlinckx, F. (2009a). Investigating latent constructs with item response models: A MATLAB IRTm toolbox. *Behavior Research Methods*, 41(4):1127–1137.

Braeken, J. and Tuerlinckx, F. (2009b). A mixed model framework for teratology studies. *Biostatistics*, 10(4):744–755.

Braeken, J., Tuerlinckx, F., and De Boeck, P. (2007). Copula functions for residual dependency. *Psychometrika*, 72(3):393–411.

Central Intelligence Agency (2008). The World Factbook. https://www.cia.gov/library/publications/the-world-factbook/index.html.

Chambers, J., Cleveland, W., Kleiner, B., and Tukey, P. (1983). *Graphical Methods for Data Analysis*. Wadsworth, Belmont, CA.

Chen, S. X. and Huang, T.-M. (2007). Nonparametric estimation of copula functions for dependence modelling. *The Canadian Journal of Statistics*, 35(2):265–282.

Chernozhukov, V., Fernández-Val, I., and Galichon, A. (2009). Improving point and interval estimators of monotone functions by rearrangement. *Biometrika*, 96:559–575.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). *Copula Methods in Finance*. John Wiley and Sons, Chichester, England.

Cherubini, U., Mulinacci, P., Gobbi, F., and Romagnoli, S. (2011). *Dynamic Copula Methods in Finance*. The Wiley Finance Series. John Wiley & Sons.

Colangelo, A. (2008). A study on LTD and RTI positive dependence orderings. *Statistics & Probability Letters*, 78:2222–2229.

Colangelo, A., Scarsini, M., , and Shaked, M. (2006). Some positive dependence stochastic orders. *Journal of Multivariate Analysis*, 97:46–78.

Colangelo, A., Scarsini, M., and Shaked, M. (2005). Some notions of multivariate positive dependence. *Insurance: Mathematics and Economics*, 37:13–26.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. *Acad. Roy. Belg. Bull. Cl. Sci.*, 65:274–292.

Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2005). *Actuarial theory for dependent risks: measures, orders and models*. Actuarial theory for dependent risks: measures, orders and models. Wiley.

Denuit, M. and Scaillet, O. (2004). Nonparametric tests for positive quadrant dependence. *Journal of Financial Econometrics*, 2(3):422–450.

Dette, H., Neumeyer, N., and Pilz, K. (2006). A simple nonparametric estimator of a strictly monotone regression function. *Bernoulli*, 12:469–490.

Dette, H. and Volgushev, S. (2008). Non-crossing non-parametric estimates of quantile curves. *Journal of the Royal Statistical Society, Series B*, 70:609–627.

Ebrahimi, N. (1993). Estimating a bivariate survial function and its marginals under positive quadrant dependence. *Journal of Statistical Computation and Simulation*, 47(1):25–35.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, Germany.

Esary, J. D. and Proschan, F. (1972). Relationships among some concepts of bivariate dependence. *The Annals of Mathematical Statistics*, 43(2):651–655.

Fan, J. and Gijbels, I. (1996). *Local polynomial modelling and its applications*. CRC Press, London, UK.

Fermanian, J.-D., Radulović, D., and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, 10(5):847–860.

Fils-Villetard, A., Guillou, A., and Segers, J. (2008). Projection estimators of Pickands dependence function. *The Canadian Journal of Statistics*, 36(3):369–382.

Fougères, A.-L. (1997). Estimation de densités unimodal. *The Canadian Journal of Statistics*, 25:375–387.

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using Copulas. *North American Actuarial Journal*, 2(1):1–25.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika*, 74(3):549–555.

Genest, C., Ghoudi, K., and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82:543–552.

Genest, C., Kojadinovic, I., Nešlehová, J., and Yan, J. (2011). A goodness-of-fit test for bivariate extreme-value copulas. *Bernoulli*, 17(1):253–275.

Genest, C., Masiello, E., and Tribouley, K. (2009a). Estimating copula densities through wavelets. *Insurance: Mathematics and Economics*, 44(2):170–181.

Genest, C., Quessy, J.-F., and Rémillard, B. (2006). Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scandinavian Journal of Statistics*, 33(2):337–366.

Genest, C. and Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 44(6):1096–1127.

Genest, C., Rémillard, B., and Beaudoin, D. (2009b). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2):199–213.

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate archimedean copulas. *Journal of the American Statistical Association*, 88(423):1034–1043.

Gijbels, I., Hall, P., Jones, M. C., and Koch, I. (2000). Tests for monotonicity of a regression mean with guaranteed level. *Biometrika*, 87(3):663–673.

Gijbels, I. and Mielniczuk, J. (1990). Estimating the density of a copula function. *Communications in Statistics – Theory and Methods*, 19(2):445–464.

Gijbels, I., Omelka, M., and Sznajder, D. (2010). Positive quadrant dependence tests for copulas. *The Canadian Journal of Statistics*, 38(4):555–581.

Gijbels, I. and Sznajder, D. (2011a). Positive quadrant dependence testing and constrained copula estimation. *Submitted.*

Gijbels, I. and Sznajder, D. (2011b). Testing tail monotonicity by constrained copula estimation. *Submitted.*

Gijbels, I. and Sznajder, D. (2011c). Testing stochastic monotonicity by constrained copula estimation. *Manuscript.*

Gijbels, I., Veraverbeke, N., and Omelka, M. (2011). Conditional copulas, association measures and their applications. *Computational Statistics & Data Analysis*, 55(5):1919–1932.

Hafner, C. M. and Manner, H. (2010). Dynamic stochastic copula models: estimation, inference and applications. *Journal of Applied Econometrics.*

Janic-Wróblewska, A., Kallenberg, W. C. M., and Ledwina, T. (2004). Detecting positive quadrant dependence and positive function dependence. *Insurance: Mathematics and Economics*, 34:467–487.

Kaas, R., Goovaerts, M., Dhaene, J., and Denuit, M. (2004). *Modern actuarial risk theory.* Kluwer Academic.

Kochar, S. C. and Gupta, R. P. (1987). Competitors of the kendall-tau test for testing independence against positive quadrant dependence. *Biometrika*, 74(3):664–666.

Kojadinovic, I. and Yan, J. (2010). Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software*, 34(9):1–20.

Kojadinovic, I. and Yan, J. (2011). A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. *Statistics and Computing*, 21:17–30.

Lee, S., Linton, O., and Whang, Y.-J. (2009). Testing for stochastic monotonicity. *Econometrica*, 77(2):585–602.

Lehmann, E. L. (1959). *Testing Statistical Hypothesis*. Wiley, New York.

Lehmann, E. L. (1966). Some concepts of dependence. *The Annals of Mathematical Statistics*, 37(5):1137–1153.

Lieb, E. H. and Loss, M. (2001). *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Rhode Island, USA, 2nd edition.

Mardia, K. V. (1970). A translation family of bivariate distributions and frechet bounds. *Sankhyā: The Indian Journal of Statistics, Series A*, 32(1):119–122.

McNeil, A. J. (1997). Estimating the tails of loss severity distributions using extreme value theory. *Astin Bulletin*, 27(1):117–137.

Meyer, M. C. (2008). An algorithm for projections onto convex cones with applications in statistical modelling. *Manuscript*.

Minicozzi, A. (2003). Estimation of sons intergenerational earnings mobility in the presence of censoring. *Journal of Applied Econometrics*, 18:291–314.

Nelsen, R. B. (2006). *An Introduction to Copulas (Lecture Notes in Statistics)*. Springer, New York, USA.

Omelka, M., Gijbels, I., and Veraverbeke, N. (2009). Improved kernel estimation of copulas: Weak convergence and goodness-of-fit testing. *The Annals of Statistics*, 37(5B):3023–3058.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Scaillet, O. (2004). Nonparametric estimation and sensitivity analysis of expected shortfall. *Mathematical Finance*, 14(1):115–129.

Scaillet, O. (2005). A Kolmogorov-Smirnov type test for Positive Quadrant Dependence. *The Canadian Journal of Statistics*, 33:415–427.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

Solon, G. (1992). Intergenerational income mobility in the united states. *American Economic Review*, 82:393–408.

Tukey, J. (1958). A problem of Berkson, and minimum variance orderly estimators. *The Annals of Mathematical Statistics*, 29:588–592.

Veraverbeke, N., Omelka, M., and Gijbels, I. (2011). Estimation of a conditional copula and association measures. *Scandinavian Journal of Statistics*. Preprint.

Wand, M. and Jones, M. (1995). *Kernel Smoothing*. Chapman & Hall.

Xu, Q. (2010). *Archimedean-Copula-Based Models in Financial Risk Management*. Lambert Academic Publishing.

Yan, J. (2007). Enjoy the joy of copulas: With a package copula. *Journal of Statistical Software*, 21(4):1–21.